Conference of PhD Students in Computer Science

Volume of extended abstracts

CS^2

Organized by the Institute of Informatics of the University of Szeged
and the Regional Committees of the Hungarian Academy of Sciences

June 28 – June 30, 2012
Szeged, Hungary

HUNGARY'S RENEWAL
The project is supported by the European Union and co-financed by the European Social Fund.
NEW SZÉCHENYI PLAN

Scientific Committee:

János Csirik (Chair, SZTE)
András Benczúr (ELTE)
Hasszan Charaf (BME)
Tibor Csendes (SZTE)
László Cser (BCE)
János Demetrovics (ELTE)
József Dombi (SZTE)
Zoltán Fülöp (SZTE)
Aurél Galántai (ÓE)
Tibor Gyimóthy (SZTE)
Zoltán Horváth (ELTE)
Csanád Imreh (SZTE)
Zoltán Kató (SZTE)
Zoltán Kása (Sapientia EMTE)
László Keviczky (SZIE)
János Kormos (DE)
László Kozma (ELTE)
János Levendovszky (BME)
Eörs Máté (SZTE)
Attila Pethő (DE)
András Recski (BME)
Lajos Rónyai (SZTAKI)
Endre Selényi (BME)
Tamás Szirányi (SZTAKI)
Péter Szolgay (PPKE)
Tibor Tóth (ME)

Organizing Committee:

Kálmán Palágyi, Rudolf Ferenc, Balázs Bánhelyi, Tamás Gergely, Zoltán Kincses

Address of the Organizing Committee

c/o. Kálmán Palágyi
University of Szeged, Institute of Informatics
H-6701 Szeged, P.O. Box 652, Hungary
Phone: +36 62 546 396, Fax: +36 62 546 397
E-mail: `cscs@inf.u-szeged.hu`
URL: `http://www.inf.u-szeged.hu/~cscs/`

Sponsors

TÁMOP-4.2.2/B-10/1-2010-0012 program of the Hungarian National Development Agency
University of Szeged, Institute of Informatics
Regional Committees of the Hungarian Academy of Sciences
Polygon Publisher

Preface

This conference is the eighth in a series. The organizers have tried to get together those PhD students who work on any fields of computer science and its applications to help them possibly in writing their first abstract and paper, and may be to give their first scientific talk. As far as we know, this is one of the few such conferences. The aims of the scientific meeting were determined on the council meeting of the Hungarian PhD Schools in Informatics: it should

- provide a forum for PhD students in computer science to discuss their ideas and research results,
- give a possibility to have constructive criticism before they present the results in professional conferences,
- promote the publication of their results in the form of fully refereed journal articles, and finally
- promote hopefully fruitful research collaboration among the participants.

The best talks will be awarded with the help of our sponsors. The papers emerging from the presented talks will be forwarded to the journals of Acta Cybernetica and Infocommunications Journal. The deadline for the submission of the papers is the middle of September 2012. The manuscripts will be forwarded to the proper journals.

Although we did not advertise it on the web, a high number of good quality abstracts have been submitted. If you encounter any problems during the meeting, please do not hesitate to contact one of the Organizing Committee members. The organizers hope that the conference will be a valuable contribution to the research of the participants, and wish a pleasant stay in Szeged.

Szeged, June 2012

Kálmán Palágyi
Rudolf Ferenc
Balázs Bánhelyi
Tamás Gergely
Zoltán Kincses

Contents

Program

Thursday, June 28

08:00 – 08:45	Registration
08:45 – 09:00	Opening
09:00 – 10:40	Talks – Mathematical problems (5x20 minutes)
10:40 – 11:00	Break
11:00 – 11:50	Plenary talk
11:50 – 12:00	Break
12:00 – 12:40	Talks – ERP (2x20 minutes)
12:40 – 14:10	Lunch
14:10 – 15:30	Talks – Recognition (4x20 minutes)
15:30 – 15:50	Break
15:50 – 17:30	Talks – Software development (5x20 minutes)
19:00	Reception

Friday, June 29

09:00 – 10:40	Talks – Data mining (5x20 minutes)
10:40 – 11:00	Break
11:00 – 11:50	Plenary talk
11:50 – 12:00	Break
12:00 – 12:40	Talks – Interpolation (2x20 minutes)
12:40 – 14:00	Lunch
14:00 – 15:20	Talks – Optimization (4x20 minutes)
15:20 – 15:40	Break
15:40 – 17:00	Talks – WEB and mobile (4x20 minutes)
18:00	Bus trip, Museum, Gala Dinner

Saturday, June 30

09:00 – 10:00	Talks – Image processing I. (3x20 minutes)
10:00 – 10:15	Break
10:15 – 11:15	Talks – Image processing II. (3x20 minutes)
11:15 – 11:35	Break
11:35 – 12:55	Talks – Image processing III. (4x20 minutes)
12:55 – 14:00	Lunch
14:00 – 15:00	Talks – Compilers (3x20 minutes)
15:00 – 15:20	Closing

Detailed Program

Thursday, June 28

8:00	**Registration**
8:45	**Opening**
Section	**Mathematical problems**
9:00	E. Antal, E. Griechisch: *Symbolic Tools for Reformulation of Unconstrained Nonlinear Optimization Problems*
9:20	B. L. Lévai: *Verified localization of trajectories with prescribed behaviour in the forced damped pendulum*
9:40	A. Bóta, M. Krész, A. Pluhár: *Efficient simulation of infection processes*
10:00	L. Hegedüs: *On Some Subsystems of Interval-Valued Logic*
10:20	L. Hegedüs, B. Nagy: *Periodicity of Circular Words*
10:40	Break
Section	**Plenary talk**
11:00	András Kornai: *Language death in the digital age*
11:50	Break
Section	**ERP**
12:00	I. Orosz, T. Orosz: *Change management applying comparison of different versions of Dynamics AX*
12:20	A. Selmeci, T. G. Orosz: *Workflow processing using SAP Objects*
12:40	Lunch
Section	**Recognition**
14:10	A. Mosavi, M. Hoffmann, A. S. Milani: *Adapting the Reactive Search Optimization and Visualization Algorithms for Multiobjective Optimization Problems; Application to Geometry*
14:30	G. Divéki: *An online 1-dimensional clustering problem with variable sized clusters*
14:50	E. Griechisch, J. Csirik: *Acceleration-based Online Signature Verification*
15:10	J. Németh: *Online Signature Feature Extraction from Video*
15:30	Break
Section	**Software development**
15:50	Zs. Borsi: *Describing interaction relationships in SMV language*
16:10	A. Góbi, Z. Szűgyi, T. Kozsik: *Embedding a Query Language into C++*
16:30	P. Hegedűs: *A Probabilistic Software Quality Model for C# – an Industrial Case Study*
16:50	Gy. Orbán: *Analysing Synchronization Contract Support in Modern Software Development Methodologies*
17:10	M. Simon: *AST-based source code analysis and visualization*
19:00	Reception

Friday, June 29

Section	**Data mining**
9:00	I. Póser: *Non-destructive Examination of Tree Trunks with the Help of National Instruments Devices*
9:20	A. Fülöp: *A Biological Driven Method for Correlation Clustering*
9:40	Gy. Móra: *Approximate dictionary matching for biomedical information extraction*
10:00	E. Földi, A. Mosavi, A. Delavar, K. N. Hewage, A. S. Milani, A. A. Moussavi, M. Yeheyis: *Reconsidering the Multiple Criteria Decision Making Problems of Construction Projects; Using Advanced Visualization and Data Mining Tools*
10:20	I. Nagy T.: *Wikipedia-based methods to identify noun compounds in running texts*
10:40	Break
Section	**Plenary talk**
11:00	Horst R. Thieme: *Apparent paradoxes in disease models with horizontal and vertical transmission*
11:50	Break
Section	**Interpolation**
12:00	P. Kovács: *Comparing wavelets and rational methods in ECG signal processing*
12:20	G. Valasek, L. Sallai, A. Jámbori, J. Horváth: *Geometric Newton-Raphson Methods for Plane Curves*
12:40	Lunch
Section	**Optimization**
14:00	V. Árgilán, J. Balogh, B. Dávid, M. Krész: *Heuristic methods for the bus rescheduling problem*
14:20	L. Lócsi, F. Schipp: *Analysis of a hyperbolic optimization method*
14:40	A. Tormási, L. T. Kóczy: *Efficiency and Accuracy Analysis of a Fuzzy Single-Stroke Character Recognizer with Various Rectangle Fuzzy Grids*
15:00	Sz. Lefkovits: *Hybrid Face Detector Based on Boosted Classifiers*
15:20	Break
Section	**Web and mobile**
15:40	E. Lázár: *PhP + MySQL questions and possible answers. Questions and possible solutions by programming a questionnary-make software.*
16:00	R. Hodován: *The Security History of the WebKit Browser Engine*
16:20	B. Keresztury, L. Cser: *Introduction to an automatized cheating prevention system at computerized open-book lab exams*
16:40	K. Pándi, H. Charaf Dr: *Synergy between mobile terminal and cloud computing*
18:00	Bus trip, Museum, Gala Dinner

Saturday, June 30

Section	**Image processing I.**
9:00	P. Bodnár: *Barcode detection with uniform partitioning and morphological operations*
9:20	T. Dobján, G. Németh: *Spectrum skeletonization: a new method for acoustic sign feature extraction*
9:40	G. Gulyás, J. Dombi: *The vagueness measure: a new interpretation and an application to image thresholding*
10:00	Break
Section	**Image processing II.**
10:15	Z. Ozsvár, P. Balázs: *Empirical Studies of Reconstructing hv-Convex Binary Matrices from Horizontal and Vertical Projections*
10:35	M. Gara and P. Balázs: *Artificial Intelligence Methods in Discrete Tomography*
10:55	N. Hantos, P. Balázs, K. Palágyi: *Binary Tomography Using Two Projections and Morphological Skeleton*
11:15	Break
Section	**Image processing III.**
11:35	T. S. Tasi, P. Balázs: *Extracting Geometrical Features of Discrete Images from Their Projections*
11:55	L. Varga, P. Balázs, A. Nagy: *A Novel Optimization-Based Reconstruction Algorithm for Multivalued Discrete Tomography*
12:15	B. Fodor: *Environment Matting by Sparse Recovery*
12:35	Á. Fazekas, H. Daimon, H. Matsuda, L. Tóth: *Computation of Background for Photoemission Electron Microscope Image Using FPGA*
12:55	Lunch
Section	**Compilers**
14:00	A. Antyipin, A. Góbi, T. Kozsik: *Low Level Conditional Move Optimization (Intel Sandy Bridge and AMD K12)*
14:20	Zs. Paróczi: *x86 instruction reordering for code compression*
14:40	T. Szirbucz: *Ideas for Improving Efficiency of Procedural Abstraction*
15:00	Closing

Symbolic Tools for Reformulation of Unconstrained Nonlinear Optimization Problems

Elvira Antal and Erika Griechisch

Formalization decisions in mathematical programming could significantly influence the complexity of the problem, and so determine the computational effort of the applied solver methods. Consequently, there is some potential in reformulation to improve the efficiency by producing easier to solve problem forms. This kind of preprocessing, working on the formula of the problem is usually done by hand in the stage of stating the problem, but that is not the only possibility. There are several techniques, that do automatic manipulations on optimization problems in order to increase the efficiency of the solver. However, such a usage of symbolic computation is applied better just for linear and integer programming. The quick evolution and the widespread use of computer algebra systems in the last few years motivated us to use symbolic computation techniques also in the field of global optimization.

We present an algorithm, what is potentially able to simplify an unconstrained nonlinear optimization problem by transforming it to an alternative problem, what can be solved with less computation. Redundant variables also can be eliminated, while the simplified task remains equivalent to the original, in the sense, that a conversion between the solutions of the two forms is possible. Our implementation based on the method of Csendes and Rapcsák [2, 3], and realized in a popular computer algebra system.

This preprocessing method seems to be unusual in nonlinear optimization, but the results for custom made and standard global optimization test problems show, that the solution is favorable in many cases.

References

[1] Antal, E., T. Csendes, and J. Virágh. Nonlinear Transformations for the Simplification of Unconstrained Nonlinear Optimization Problems, submitted. `http://www.inf.u-szeged.hu/~antale/en/research/Antale_Opkut2011.pdf`

[2] Csendes, T., T. Rapcsák. Nonlinear Coordinate Transformations for Unconstrained Optimization. I. Basic Transformations, *J. of Global Optimization* 3(1993) 213–221

[3] Rapcsák, T., T. Csendes. Nonlinear Coordinate Transformations for Unconstrained Optimization. II. Theoretical Background., *J. of Global Optimization* 3(1993) 359–375

Low Level Conditional Move Optimization (Intel Sandy Bridge and AMD K12)

Artyom Antyipin, Attila Góbi, and Tamás Kozsik

Although high level optimizations are becoming more and more sophisticated, the importance of low level optimizations should not be underestimated either. Sensible performance improvements can be achieved simply by reordering instructions or using an alternative, but equivalent instruction sequence. Besides all the enhancements implemented in modern processors they remain compatible to their ancestors – new instruction sets are introduced, but none of the existent instructions become unsupported. Due to the changes in the inner architecture of the processors, some optimization techniques may become more or less effective. Existing techniques need, from time to time, be reconsidered, and new techniques, targeting these modern architectures, may emerge.

This paper introduces a novel approach to the optimization of the conditional move operations, using the well known instruction introduced in the sixth generation Intel x86 microarchitecture: `CMOVcc`. In spite of being around for a long time, it was never used actively – frequently used code generators like gcc or cl.exe also avoid it. This decision can be explained by the fact that the impact of this instruction on the efficiency of the generated code is not obvious – inappropriate use may result in sensible performance regression. Furthermore, in most cases the replacement of this instruction with a corresponding conditional jump and unconditional move yields performance improvement on older processors.

Being forgotten and avoided, `CMOVcc` not only remained supported by all modern processors, but became recommended both by AMD [1] and Intel [2,3] as an alternative to the sequence of a conditional jump and an unconditional move. One of the main reasons of this recommendation is that the usage of the conditional move avoids any possibility of branch mis-prediction (which is inescapable when a conditional jump depends on random data), and any further performance penalties caused by the mis-prediction.

In this paper we analyze the usage of `CMOVcc` in different contexts on various modern AMD (K12 family) and Intel (Sandy Bridge family) processors, and, based on these results, propose a technique to automatically decide whether conditional move or the sequence of a conditional jump and an unconditional move should be performed in a given situation.

Acknowledgements

Supported by the European Union and co-financed by the European Social Fund (grant agreement no. TAMOP 4.2.1./B-09/1/KMR-2010-0003).

References

[1] Advanced Micro Devices, Inc. Software optimization guide for AMD family 10h and 12h processors. `http://support.amd.com/us/Processor_TechDocs/40546.pdf`.

[2] Intel Corporation. Intel® 64 and IA-32 architectures optimization reference manual. `http://www.intel.com/content/dam/doc/manual/64-ia-32-architectures-optimization-manual.pdf`.

[3] Intel Corporation. Intel® 64 and IA-32 architectures software developer's manual. `http://www.intel.com/content/dam/doc/manual/64-ia-32-architectures-software-developer-vol-1-2a-2b-3a-3b-manual.pdf`.

Heuristic methods for the bus rescheduling problem

Viktor Árgilán, János Balogh, Balázs Dávid, Miklós Krész

Recovery from disruptions and the rescheduling of vehicles is a fairly new research field in the optimization of public transportation [1,2,3]. Vehicle disruptions can occur as a result of a variety of reasons: accidents, roadblocks, or simply because of the delay of one or more trips. Since disruptions happen on a daily basis, transportation companies have to consider a proper method to deal with this problem.

The typical way for companies to address the delay of one trip is by using a backup vehicle from the depot. However, the above approach might not always prove to be the most efficient solution, and dealing with several delayed trips simultaneously would require more backup vehicles present in the depot. Different modes of public transportation require various approaches for the solution of this problem. First, we give an overview of the models and methods in literature for the different types of vehicle rescheduling problems (aircraft, rolling stock, bus). Based on these methods, a mathematical model is developed for the multiple depot bus rescheduling problem.

However, certain aspects have to be considered when applying a solution method in practice. Most importantly, a practical situation needs a real time solution for the problem in order to guarantee efficiency in operations management. Fast heuristic approaches can be applied to give suggestions and help in supporting the decision itself. The two most important features of these methods are:

i) give a well-structured, realizable solution for the problem,

ii) solutions for the problem are expected to be produced in real time

To address the above requirements, we developed several types of solution algorithms for the problem which can help a transportation company in their decision support process.

Acknowledgements

This work was supported by Gyula Juhász Faculty of Education, University of Szeged (project number CS-004/2012.V.3).

References

[1] J. Clausen, A. Larsen, J. Larsen, N.J. Rezanova: Disruption Management in the Airline Industry - Concepts, Models and Methods. Computers & Operations Research, Vol. 37., 809-821., 2010.

[2] J. Jespersen-Groth, D. Potthoff, J. Clausen, D. Huisman, L. Kroon, G. Maroti, M.N. Nielsen: Disruption Management in Passenger Railway Transportation. Lecture Notes in Computer Science, Vol. 5868/2009, 399-421., 2009.

[3] J-Q. Li, D. Borenstein, P. B. Mirchandani: A decision support system for the single-depot vehicle rescheduling problem. Computers & Operations Research Vol. 34., 1008-1032., 2007.

Barcode detection with uniform partitioning and morphological operations

Péter Bodnár

Visual codes referred as barcodes have two major groups, 1- and 2-dimensional barcodes. 1D barcodes consist of a well-defined group of parallel lines aiming easy automatic identification of carried data with endpoint devices such as PoS terminals, smartphones, or computers. Most two-dimensional barcodes serve the same purpose as their one-dimensional ancestors. They usually consist of simple picture elements (e.g. rectangles, dots, hexagons) organized in a 2D pattern layout.

Barcode detection methods have two main objectives, speed and accuracy. On smartphones, fast detection of barcodes is desirable, but accuracy is not so critical since the user can easily reposition the camera and repeat the scan. For industrial environment, accuracy and speed are desired properties for detection.

For 1D barcodes, the basic approach for detection is scanning only one, or just a couple of lines of the whole image. This method is common at hand-held PoS laser scanners or smartphone applications. Scanned lines form an 1D intensity profile, and barcode-detector algorithms [1, 2, 3] work on these profiles to find an ideal binary function that represents the original encoded data. The main idea is to find peak locations in blurry barcode models, then thresholding the intensity profile adaptively to produce binary values. For 2D barcodes, most approaches involve the extraction of texture-like properties and detection of properties that refer to code-like appearance. Also, there are approaches for detection of specific types of codes with heavy perspective distortion, motion blur [4], and noise [5].

The MIN-MAX method [6, 7, 8] uses the difference image of the dilated and eroded original image. White blobs on these images show the possible barcode locations. Further processing, like segmentation and filtering of small blobs are required on these difference images. This approach can be used on both 1D and 2D barcodes with minor modifications. Methods based on wavelet transformation [9] look at images for barcode-like appearance by a cascaded set of weak classifiers. Each classifier working in the wavelet domain narrows down the possible set of barcodes, decreasing the number of false positives while trying to keep the best possible accuracy. Variants of Hough transformation [10] detect barcodes by working on the edge map of the image. The two most common methods are standard and probabilistic Hough transformation. Both transform edge points into Hough space first, and make decisions of line locations. Our proposed method examines the image in small, disjunct or overlapping tiles, and make local measurements. We use distance transformation [11], clustering and morphological operations for these measurements.

References

[1] Timothy R. Tuinstra. *Reading Barcodes from Digital Imagery*. PhD thesis, Cedarville University, 2006.

[2] Eugene Joseph and Theo Pavlidis. Bar code waveform recognition using peak locations. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 16(6):630–640, jun 1994.

[3] Péter Bodnár and László G. Nyúl. Fast barcode detection with extended scanline analysis. In *Advanced Concepts for Intelligent Vision Systems, 2012., Proceedings of Conference*, 2012. under review.

[4] Chung-Hua Chu, De-Nian Yang, Ya-Lan Pan, and Ming-Syan Chen. Stabilization and extraction of 2d barcodes for camera phones. *Multimedia Systems*, 17:113–133, 2011. 10.1007/s00530-010-0206-9.

[5] Jong-Eun Ha. A new method for detecting data matrix under similarity transform for machine vision applications. *International Journal of Control, Automation and Systems*, 9:737–741, 2011. 10.1007/s12555-011-0415-9.

[6] Péter Bodnár and László G. Nyúl. Efficient barcode detection with texture analysis. In *Signal Processing, Pattern Recognition and Applications, 2012., Proceedings of Conference*, 2012. accepted for publication.

[7] Daw-Tung Lin, Min-Chueh Lin, and Kai-Yung Huang. Real-time automatic recognition of omnidirectional multiple barcodes and dsp implementation. *Machine Vision and Applications*, 22:409–419, 2011. 10.1007/s00138-010-0299-3.

[8] Daw-Tung Lin and Chin-Lin Lin. Multi-symbology and multiple 1d/2d barcodes extraction framework. In Kuo-Tien Lee, Wen-Hsiang Tsai, Hong-Yuan Liao, Tsuhan Chen, Jun-Wei Hsieh, and Chien-Cheng Tseng, editors, *Advances in Multimedia Modeling*, volume 6524 of *Lecture Notes in Computer Science*, pages 401–410. Springer Berlin / Heidelberg, 2011. 10.1007/978-3-642-17829-0-38.

[9] R. Oktem. Bar code localization in wavelet domain by using binary morphology. In *Signal Processing and Communications Applications Conference, 2004. Proceedings of the IEEE 12th*, pages 499 – 501, april 2004.

[10] D.H. Ballard. Generalizing the hough transform to detect arbitrary shapes. *Pattern Recognition*, 13(2):111 – 122, 1981.

[11] Pedro F. Felzenszwalb and Daniel P. Huttenlocher. Distance transforms of sampled functions. Technical report, Cornell Computing and Information Science, 2004.

Describing interaction relationships in SMV language

Zsolt Borsi

During the phases of the software development process Unified Modeling Language (UML) helps in managing the complexity of problems by separately describing different aspects of the system under development. The created intermediate artifacts are mainly UML diagrams representing a particular view of the system. Since UML models represent specification documents, they provide ideal basis for validation and verification. Model checking is a successful method to prove that a given system meets its specification [3, 4]. It has been found especially useful when the correctness of the system requires a formal approach, for instance when the system under consideration is a safety-critical system.

In this paper the translation of a particular scenario-based model into SMV language (the input language of the NuSMV model-checker tool [5]) will be investigated. With the introduction of Interaction Overview Diagrams (IODs) in UML2.0, relationships between interaction diagrams can be explicitly defined in a standard way. IODs are a graphical representation of relationships between UML interaction diagrams and the control flow passing between them, where each node in the activity graph is a reference to an IOD. This paper takes into account additional relationships, which are not available in IODs. These extension version of Interaction Overview Diagram (namely EIOD) introduced by Whittle [1] includes interruption, continuation, concurrency and the notion of negative scanario. Whittle introduced a way of structured specification as well. A use case chart specifies the scenarios for a system's use cases in such a way that each of the EIOD nodes is refined by a set of use case nodes at the lower level of the use case chart. In fact, in this paper the description of a use case chart in SMV language will be presented. The rationale behind that idea is that the module concept of SMV provides a mean for describing a hierarchical construct and SMV modules (like scenarios defined by interaction diagrams) operate parallel to each other.

Acknowledgements

This work was supported by TÁMOP-4.2.1/B-09/1/KMR-2010-0003

References

[1] J. Whittle, P.K. Jayaraman. Synthesizing hierarchical state machines from expressive scenario descriptions, *ACM Transactions on Software Engineering and Methodology (TOSEM), v.19 n.3, pp.1-45, 2010.*

[2] E.M. Clark, W. Heinle. Modular Translation of Statecharts to SMV, *Technical Report CMU-CS-00-XXX, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213., 2000.*

[3] E.M. Clark, O. Grumberg, D.A. Peled. Model Checking, *The MIT Press, Cambridge, MA, 2000.*

[4] O. Kupferman, M.Y. Vardi, P. Wolper. Module checking, *Information and Computation, 164(2), pp.322-344, 2001.*

[5] NuSMV Model Checker Home Page, *http://nusmv.fbk.eu*

Efficient simulation of infection processes

András Bóta, Miklós Krész, András Pluhár

Studying virus marketing Domingos and Richardson [1] introduced the so-called Independent Cascade model. Kleinberg, Kempe and Tardos have proven in [2] that this model is an equivalent form of Granovetter's Linear Threshold model [3]. In that model a small set, the initial adopters (infected vertices) infect neighboring vertices with prescribed probabilities in stages. This model has been extended and applied to many other real-world examples with considerable success [4].

The computation of a single infection process may not be difficult, but very often we are faced with the task of predicting the possibility of nodes (or the network as a whole) getting infected. This task can be solved by running the infection process multiple times resulting in an *a posteriori* probability distribution over the network. Due to the nature of infection dynamics the complexity of this simulation is demanding.

Another way of solving this task is to design an algorithm that is able to compute the a posteriori distribution directly, omitting the simulation. Unfortunately, this task is also very difficult. In this talk we will describe alternative ways of simulating infection processes with improved performance. We will also describe an algorithm, that is able to compute the desired distribution directly. Finally we will compare the results of these methods considering both accuracy and time complexity.

Acknowledgements

The authors were supported by the Project named "TÁMOP-4.2.1/B-09/1/ KONV-2010-0005 - Creating the Center of Excellence at the University of Szeged" also supported by the European Union and co-financed by the European Regional Fund.

The second author was partially supported by Gyula Juhász Faculty of Education, University of Szeged.

The third author was partially supported by the grant OTKA K76099.

References

[1] P. Domingos and M. Richardson. Mining the network value of customers. In Proc. 7th Intl. Conf. on Knowledge Discovery and Data Mining, (2001) 57–66.

[2] D. Kempe, J. Kleinberg, and É. Tardos. Maximizing the spread of influence through a social network. In Proc. KDD, 2003.

[3] M. Granovetter. Threshold models of collective behavior. American Journal of Sociology 83 (1978) 1420–1443.

[4] A. Csernenszky, Gy. Kovács, M. Krész, A. Pluhár and T. Tóth. The use of infection models in accounting and crediting. Proc. of the Challenges for Analysis of the Economy, the Business, and Social Progress International Scientific Conference. Szeged, November 19–21, 2009.

An online 1-dimensional clustering problem with variable sized clusters

Gabriella Divéki

In the online clustering problems, the classification of points into sets (called clusters) is done in an online fashion. Points arrive one by one at arbitrary locations, to be assigned to clusters at the time of arrival without any information about the further points. A point can be assigned to an existing cluster, or a new cluster can be opened for it. Existing clusters cannot be merged or split. We study one-dimensional variants. The cost of a cluster is the sum of a fixed setup cost scaled to 1 and the square of the diameter of the cluster. The goal is to minimize the sum of costs of the clusters used by the algorithm.

We study two variants, both maintaining the properties that a point which was assigned to a given cluster must remain assigned to this cluster, and clusters cannot be merged. In the strict variant, the size and the exact location of the cluster must be fixed when it is initialized. In the flexible variant, the algorithm can shift the cluster or expand it, as long as it contains all points assigned to it. We study the online and the semi-online versions of the above two variants.

In [1] the one-dimensional variant of our problem is examined, where there is no restriction on the length of a cluster, and the cost of a cluster is the sum of a fixed setup cost and its diameter. Both the strict and the flexible model have been investigated and an intermediate model, where the diameter is fixed in advance but the exact location can be modified is also studied.

In [2] the two-dimensional problem is considered using the l_∞ norm instead of the l_2 norm. Thus, "balls" are actually squares. The cost of each cluster is a sum of a constant setup cost scaled to 1 and the area of the cluster. The goal is to minimize the total cost of the clusters.

We present the first online algorithms for the solution of the problem. We present algorithms for the strict and the flexible variant. We also give lower bounds on the possible competitive ratio in the case of both of the variants.

References

[1] J. Csirik, L. Epstein, Cs. Imreh and A. Levin, Online Clustering with Variable Sized Clusters, *In Proceedings of MFCS 2010, LNCS 6281* 282–293, 2010.

[2] G. Divéki and Cs. Imreh, An online 2-dimensional clustering problem with variable sized clusters, *Submitted for publication, 2011*.

Spectrum skeletonization: a new method for acoustic sign feature extraction

Tibor Dobján and Gábor Németh

Nowadays, the noisy acoustic signals investigating by machine learning algorithms is a frequently researched topic in the signal processing. They play an important role in speech recognition algorithms and industrial applications as well. To find the appropriate parameters is a keypoint of the learning process of machine learning. The parameter sets are optimal if they produce different results for objects having different quality and similar results for those ones that are similar in quality. Methods, that are used in practice, search a threshold hit in the given time series or provide some parameters extracted from the auto power spectrum density (APSD) function (e.g., the Root Mean Square Rate calculation). In present work we propose a novel method to analyze the APSD function. The benefits of the skeletonization can be useful to determine relevant frequency components in the APSD function [1].

Skeleton is a shape descriptor which summarizes the general form of objects [2]. Although the power spectrum is a vector of density values corresponding to different frequencies, we can map it into a binary image. Centerlines provide structural and geometric information about the APSD function. The relevant peaks in it are represented by skeletal branches. Removing the "main branch" of the skeleton (i.e., the lower curve segments that ensure the connectivity between the vertical side branches), we split it into disjunct trees. The frequency having the minimal density value between the skeletal trees is the proper location to separate the frequency ranges. By this way no peak are split unlike in the former methods. In Fig. 1, the vertical dark gray lines indicate the bounds of the frequency ranges, and the skeletal trees are also superimposed with black curves in the picture. Usually the skeletal trees are inside in their frequency ranges, however in some cases the lower branch of a skeletal tree may grow into the neighbour range. In order to separate the frequency ranges we have to consider the upper side branches of a skeletal tree (i.e., that represents a peak of the APSD). Furthermore, the geometry of skeletal branches in frequency range holds useful information that can be used for shape analysis.

According to our best of knowledge, there is no other acoustic verification system that uses skeletons to analyze the structure of the APSD function of a vibration sample.

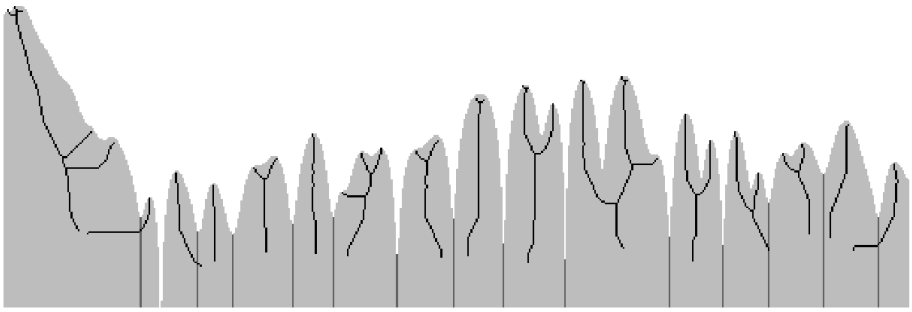


Figure 1: The APSD function of a vibration sample. Its skeletal branches are superimposed with black curves, while the frequency component bounds are indicated by dark gray vertical lines.

References

[1] T. Dobján, Sz. Pletl, T. Deák, L. Doszpod, and G. Pór: Identification of the place and materials of knocking objects in flow induced vibration, Acta Cybernetica 20 (2011) pp. 53-67.

[2] K. Siddiqi, S.M. Pizer (Eds.), Medial Representations – Mathematics, Algorithms, and Applications, Series in Computational Imaging, Springer, 2008.

Computation of Background for Photoemission Electron Microscope Image Using FPGA

Ádám Fazekas, Hiroshi Daimon, Hiroyuki Matsuda and László Tóth

The purpose of our development is to design an FPGA based hardware acceleration system that is able to be used for analyzing photoemission electron microscope (PEEM) [1, 2, 3] images or improving their quality. Even though a usual PEEM has an energy filter unit, which is able to eliminate certain disturbing signals, a post processing computation can also be useful to improve the image quality.

Here we propose an FPGA based hardware acceleration system for the computation of a certain image background component. It has uniquely designed hardware modules that perform the computations in parallel, resulting in less calculation time. The system shown here is a prototype which was only used for testing and experimental purposes.

The prototype of the background computer system provided remarkable results and important experiences which will be useful in the design of the final high performance hardware acceleration system. During the development of the prototype we realized that relevant performance enhancement requires a high-end FPGA platform which has the necessary resources to determine the background values in real-time. That device could be used as a built-in or with other words embedded unit with the measuring instrument so it is no longer necessary to use communication protocols between the PC and the hardware. The running time could be easily reduced in the future work, if the system performs the computations in more than two threads of the Shirley modules, which is limited mainly by the applied hardware resources and not by the realization of the method, therefore using a higher performance FPGA, more parallel Shirley modules could be executed simultaneously.

We have also concluded that the removal of Shirley [4] (or analogous e.g. linear) background can enhance the quality of measured images, and our method can also be applied in other fields and different kind of microscopy for different kind of background removals.

References

[1] Daimon H., Matsuda H., Tóth L.: *Stereo-PEEM for three-dimensional atomic and electronic structures of microscopic materials* in Surface Science, 601(20) (2007) pp. 4748-4758.

[2] Goto K., Matsuda H., Hashimoto M., Nojiri H., Sakai C., Matsui F., Daimon H., Tóth L., Matsushita T.: *Development of display-type ellipsoidal mesh analyzer* in e-Journal of Surface Science and Nanotechnology Vol. 9 (2011) pp. 311-314.

[3] Tóth L., Matsuda H., Matsui F., Goto K. and Daimon H.: *Details of 1 π sr wide acceptance angle electrostatic lens for electron energy and two-dimensional angular distribution analysis combined with real space imaging* in Nuclear Instruments and Methods in Physics Research A, Vol. 661, (2011) pp. 98-105.

[4] Shirley D. A: *High-Resolution X-Ray Photoemission Spectrum of the Valence Bands of Gold* in Phys. Rev. B, 55, 4709 (1972).

Environment Matting by Sparse Recovery

Bálint Fodor

Introduction

Environment matting is an image based relighting technique that aims to extract information about the light transport of a given scene exploiting the assumptions of the measurement setting. The measurement involves a digital camera and a controllable screen as the background of the scene.

While displaying special patterns in the background the camera captures the light response of the scene. Processing such measurements a background-free representation of the scene is aimed to be produced. The technique of environment compositing is to offline *re*render the scene with arbitrary background using the extracted representation.

Based on recent results in the field of compressive sensing and sparse signal recovery [1] this work solves the problem of environment matting by assuming that only few background pixels are linearly interplaying to produce a single camera pixel.

Thus a vector of corresponding pixel intensities can be thought of as a linear combination of some entries of a known signal dictionary built from vectors of corresponding pixel intensities through the displayed background patterns.

Previous Work

Whilst related works on image matting deal with basically opaque scene elements [4] the concept of environment matting intruduced by Zongker *et al.* [7] is intended to reveal the reflective and refractive nature of the scene.

Wexler *et al.* [5] formulates and solves a more general problem where the background varies due to camera movement. It involves two steps: first a merged pure background is extracted, then the 2D-2D map of the light transport is estimated. Experiments show that the method has limited accuracy especially when employing non-planar backgrounds.

Frequency and wavelet based approaches [6] are more robust but require a large number of images making the measurement last long.

Efforts on improving accuracy and time-consumption are also made [2]. In [2] the mapping between the background and the camera image is estimated by color-coding all the pixels in the backgound image and matching those of in the camera image.

The most related method of this work is attributed to Duan *et al.* [3]. A compressive sensing scheme is used to estimate the contribution coefficients of pixels by displaying patterns in a coarse-to-fine manner. It is assumed and exploited that the background pixels supporting a camera pixel (i.e. pixels that have non-zero coefficient) are spatially grouped in the background image.

Methodology

It is assumed that there is a linear relationship between the pixels of the screen and that of the camera. So the common gamma function-like intensity distortion of the devices are compensated.

The number of camera and screen pixels are denoted by d_c and d_s respectively.

A pixel value of the camera is modelled by a linear combination of at most k screen pixels:

$$c_i = \sum_{j=1}^{d_s} \alpha_j s_j, \quad |\{j : \alpha_j \neq 0\}| \leq k \tag{1}$$

where c_i is the ith pixel of the camera, s_j is the jth pixel of the screen and $0 \leq \alpha_j \leq 1$ are the screen pixel coefficients. The aim is to extract the k non-zero α_j coefficients.

Displaying M patterns (1) can be easily rewritten as $C = AS$ where C is $d_c \times M$, S is $d_s \times M$ matrices and A is a $d_c \times d_s$ sparse matrix. Each corresponding coloumns of S and C represent a pattern-response image pair. As A is the coefficient matrix it has at most k non-zero entries in each row.

Compressed sensing theory suggests employing random patterns for the problem of sparse recovery so the entries of S are choosen to be realizations of a random variable with normal distribution. Here S is called the signal dictionary. For every camera pixel a row of C belongs.

Given x a row of C the problem is to find k rows (k non-zero coefficient) in S that all together are the best approximation of x.

A is reconstructed via the Orthogonal Matching Pursuit algorithm. It is a relatively simple, greedy method and iteratively decreases the residual error by selecting the vector from the dictionary that is maximizing the dot product with the residual. The coefficients of the so far selected entries are updated in each step.

Finding the first k entries and their coefficients gives the sparse reconstruction of x. The pixels corresponding to the selected entries of S are treated as the main contributors that form x.

In this work the problem of environment matting is considerd employing a compressive sensing approach. Using OMP as the sparse solver the performance of the concept is evaluated and its competitiveness is justified.

References

[1] E.J. Candès and M.B. Wakin. An introduction to compressive sampling. *Signal Processing Magazine, IEEE*, 25(2):21–30, 2008.

[2] B. Choudhury, D. Singla, and S. Chandran. Fast color-space decomposition based environment matting. In *Proceedings of the 2008 symposium on Interactive 3D graphics and games*, page 1. ACM, 2008.

[3] Qi Duan, Jianfei Cai, Jianmin Zheng, and Weisi Lin. Fast environment matting extraction using compressive sensing. In *Multimedia and Expo (ICME), 2011 IEEE International Conference on*, pages 1 –6, july 2011.

[4] J. Wang and M.F. Cohen. Image and video matting: a survey. *Foundations and Trends® in Computer Graphics and Vision*, 3(2):97–175, 2007.

[5] Y. Wexler, A.W. Fitzgibbon, and A. Zisserman. Image-based environment matting. In *Proceedings of the 13th Eurographics workshop on Rendering*, pages 279–290. Eurographics Association, 2002.

[6] J. Zhu and Y.H. Yang. Frequency-based environment matting. In *Computer Graphics and Applications, 2004. PG 2004. Proceedings. 12th Pacific Conference on*, pages 402–410. IEEE, 2004.

[7] D.E. Zongker, D.M. Werner, B. Curless, and D.H. Salesin. Environment matting and compositing. In *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pages 205–214. ACM Press/Addison-Wesley Publishing Co., 1999.

Reconsidering the Multiple Criteria Decision Making Problems of Construction Projects; Using Advanced Visualization and Data Mining Tools

E. Foldi, A. Mosavi, A. Delavar, K. N. Hewage, A. S. Milani, A. A. Moussavi, M. Yeheyis

The construction project managers have to make tough decisions. They are considering different tools of IT and would like to invest on getting better data analysis tools for enhancing their decisions. The increasing of IT usage in construction projects has been the most effective option for improving the process of problem solving [1, 4, 6]. However making critical decisions for the complicated and multiple criteria problems of construction projects cosidering multiple coflicting criteria in which a huge amount of data involved is not a simple task to do. As the datasets of our problems, i.e. construction projects analysis dealing with waste management, productivity improvement, environmental impacts, human and IT factors, emergy based lifecycle, and process optimization, are often huge they can not easily be handled with the traditional means of data analysis. In order to better manage the data collected and make the most of our datasets a series of the advanced interactive visualization and data mining tools [5] are utilized. Using data mining tools for processing the large-scale data involved in the multiple criteria decision-making problems is previously considered in [2, 7].

Here the idea for solving the multiple criteria decision-making problems is to visually model and clarify the whole dimension of the problems. Interesting patterns are automatically extracted from our raw dataset via the data mining tools e.g. reducing the dimensionality and the dataset size, multidimensional scaling, clustering and user interaction. Additionally the advanced visual analytical interfaces are involved to support the decision maker interactively. With such tools i.e. parallel filters and clustering tasks, the decision maker can solve multiobjective optimization problems as it amends previous approaches [7]. Furthermore the animations of sweeping through data and advanced visualizations including 7D plots accomplish managers and enable them to screen the data at their consulting room making decision interactively.

The effectiveness and performance of the interactive visualizations are evaluated along with a number of study cases related to construction workers [1]. Along with our study cases the aspects of data mining, modeling, and visualization the data related to construction workers are considered and briefly presented in this article. As the main result, the achieved hidden information through visualization tools has enhanced the final decisions. Moreover the 7D plots and the option of sweeping through data and clustering [3, 5] have been found to be very useful for our challenging applications.

Keywords: Building construction workers, IT usage in construction projects, reactive business intelligence, reactive search, multiobjective optimization, multiple criteria decision making, interactive visualization, multi-dimensional plots, Grapheur, LIONsolver, 7D graphs, Clustering and sweeping through data

References

[1] K.N. Hewage, J.Y. Ruwanpura, G.F. Jergeas (2009), IT Usage in Alberta's Building Construction Projects: Current Status and Challenges, Journal of Automation in Construction. V17, pp. 940–947.

[2] A. Mosavi(2010). Multiple criteria decision-making preprocessing using data mining tools. IJCSI International Journal of Computer Science Issues, V7, Issue 2, No 1.

[3] R. Battiti, M. Brunato, Reactive Business Intelligence. From Data to Models to Insight, Reactive Search Srl, Italy, (2011).

[4] A. Mosavi, M. Azodinia, Kasun N. Hewage, Abbas S. Milani, M.Yeheyis, Reconsidering the Multiple Criteria Decision Making Problems of Construction Workers Using; Grapheur, Enginsoft newsletter, Year 8 No. 4, winter 2011.

[5] R.Battiti, M.Brunato (2010). "Grapheur: A Software Architecture for Reactive and Interactive Optimization, Proceedings Learning and Intelligent OptimizatioN LION 4, 2010, Venice, Italy.

[6] A.Mosavi, M.Azodinia, Kasun N. Hewage, Abbas S. Milani, M.Yeheyis, Reconsidering the Multiple Criteria Decision Making Problems of Construction Workers Using; Grapheur, Poster presented at EnginSoft International Conference, Verona, Italy 20-21 Oct.

[7] A.Mosavi (2010), the large-scale system of multiple criteria decision making; pre-processing, Large Scale Complex Systems Theory and Applications, V9, Part# 1.

A Biological Driven Method for Correlation Clustering

András Fülöp

Many applications benefit from partitioning data points into clusters. Correlation clustering is a clustering method which originates from document clustering. The problem was introduced by Bansel, Blum and Chawla [1] and was proved to be NP-complete in the same paper. Correlation clustering can be viewed as a graph partitioning problem in which each connection is labelled $+$ (similar) or $-$ (different) depending on the components. In correlation clustering, unlike a large portion of clustering algorithms, the number of clusters is not specified. The goal is to partition the nodes into clusters of similar nodes. There are several existing algorithms for correlation clustering (e.g. [2], [3], [4]).

In this talk we introduce a method with which the correlation clustering problem can be solved. The algorithm is based on a biological optimization method, the Ant System method introduced by Dorigo, Maniezzo, and Colorni in an early report [6], and motivated by the work of Liu and Fu who utilized a similar algorithm for clustering problems [5]. In our talk we discuss our solution and results including comparisons with other existing solutions.

References

[1] Nikhil Bansal, Avrim Blum, and Shuchi Chawla. *Correlation clustering.* Machine Learning Journal (Special Issue on Theoretical Advances in Data Clustering), 2004.

[2] Becker, Hila. *A Survey of Correlation Clustering.* COMS E6998: Advanced Topics in Computational Learning Theory, 2005.

[3] Zimek, Arthur. *Correlation Clustering.* Thesis, Ludwig-Maximilians University, Munich, 2008.

[4] Achtert, Elke, B ohm, Christian, Kriegel, Hans-Peter, Kr oger, Peer, Zimek, Arthur. *Robust, Complete, and Efficient Correlation Clustering.* SIAM Proceedings, 2007.

[5] Liu, Xiaoyong, Fu, Hiu. *An Effective Clustering Algorithm With Ant Colony.* Journal of Computers, vol. 5, no. 4, 2010.

[6] Dorigo, Maniezzo, Colorni. *Positive feedback as a search strategy.* Technical report, ipartimento di Elettronica, Politecnico di Milano, 1991.

Artificial Intelligence Methods in Discrete Tomography

Mihály Gara and Péter Balázs

Tomography is an imaging procedure to examine the internal structure of objects. The cross-section images are constructed with the aid of the object's projections. It is often necessary to minimize the number of those projections to avoid the damage or destruction of the examined object, since in most cases the projections are made by destructive rays.

Sometimes the number of available projections are so small that conventional methods cannot provide satisfactory results. In these cases *Discrete Tomograpy* can provide acceptable solutions, but it can only be used with the assumption the object is made of only a few materials, thus only a small number of intensity values appear in the reconstructed cross-section image.

Although there are a lot of discrete tomographic reconstruction algorithms, only a few papers deal with the determination of intensity values of the image, in advance. In our work we try to fill this gap by using different learning methods. During the learning and classification we used the projection values as input arguments.

In the second part of our talk we concentrate on *Binary Tomography* (a special kind of Discrete Tomography) where it is supposed that the object is composed of one material. Thus, there can be only two intensities on the cross-section image - one for the object points and one for the background. Here, we compared our earlier presented binary tomographic evolutionary reconstruction algorithm to two others. We present the details of the above-mentioned reconstruction method and our experimental results. This paper is based on our previous works [1, 2].

Acknowledgements

This work was supported by the TÁMOP-4.2.2/08/1/2008-0008 and the TÁMOP-4.2.2/B-10/1-2010-0012 program of the Hungarian National Development Agency. The work of the second author was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences and by the OTKA PD100950 grant of the Hungarian Scientific Research Fund.

References

[1] P. Balázs and M. Gara. An Evolutionary Approach for Object-Based Image Reconstruction Using Learnt Priors, *Lecture Notes in Comput. Sci.*, 5575:520–529, 2009.

[2] M. Gara, T. S. Tasi and P. Balázs. Machine Learning as a Preprocessing Phase in Discrete Tomography, *Lecture Notes in Comput. Sci.*, in Press.

Embedding a Query Language into C++

Attila Góbi, Zalán Szűgyi, and Tamás Kozsik

The integration of query languages and programming languages is an old, returning problem. In the case of SQL, traditional solutions were either based on language embedding or on representing queries as string values. The latter approach is very fragile since no compile time checks are performed. The former (relational) approach re-emerges in the object-oriented paradigm with the rise of LINQ, which appeared in .Net 3.5. LINQ is a language integrated into .Net languages providing native data query capabilities. The upcoming JavaSE 8 standard is planned to provide similar functionality, and until then Criteria Queries of JPA 2.0 can be used.

This paper addresses the same problem in C++. We develop an embedded language using the `boost::proto` library [2]. This library provides a quasi standard way to embed languages into C++. Unlike the above mentioned solutions, we are not trying to introduce a language similar to SQL or OQL, but one which is based on object comprehensions. Object comprehensions are introduced by D.K.C. Chan and P.W. Trinder [1] and are designed especially for object-oriented databases.

Due to our approach, namely the application of C++ template metaprogramming, it is possible to process a significant part of the language in compile-time. Similarly to LINQ, our query language can be compiled to work on collections as well as on databases. Furthermore, the processor of the query language is designed to be non-intrusive – an existing source code, or even a compiled object code can be integrated with the queries by providing the necessary metainformation.

Acknowledgements

Supported by the European Union and co-financed by the European Social Fund (grant agreement no. TAMOP 4.2.1./B-09/1/KMR-2010-0003).

References

[1] D. Chan, P. Trinder. Object comprehensions: A query notation for object-oriented databases. *D. Bowers (ed.) Directions in Databases, Lecture Notes in Computer Science*, vol. 826, pp. 55–72. Springer Berlin / Heidelberg (1994).

[2] E. Niebler. Proto: a compiler construction toolkit for DSELs. *Proceedings of the 2007 Symposium on Library-Centric Software Design*. pp. 42–51. ACM, New York, NY, USA (2007). `http://doi.acm.org/10.1145/1512762.1512767`

Acceleration-based Online Signature Verification

Erika Griechisch and János Csirik

Several kinds of technology are available for biometric authentication. Some of them have appeared in the last few decades - such as DNA and iris recognition -and they provide more accurate results than the earlier methods did (like fingerprint and signature recognition). Not surprisingly, these systems are harder to fool. However, a signature is still the most widely accepted method for identification (e.g. in contracts and bank transfers). For this reason, some studies tackle the problem of signature verification and examine the process in detail, as we do here.

There are two basic approaches of recognising signatures: offline and online. Offline signature recognition is based on the image of the signature, while the online case utilises data related to the dynamics of the signing process. The drawback with the offline approach is that it produces more false accept and false reject errors, but the dynamic approach requires much more sophisticated techniques. One key issue with online signature verification is to learn which features best represent the dynamics of the signing process - like velocity, acceleration and pressure.

In our view, tablets affect the natural signing process, hence we devised an online signature recording environment. This environment does not need any special surface and it contains a simple ball-point pen with a three-axis accelerometer that does not significantly affect the signature writing process. Moreover, it can be easily connected to a computer via a USB and signals can be measured using a fully open-source package called EDAQ [1, 2].

In our study, we created two databases in order to compare the signatures obtained in two different time periods and to analyse the relevance of training data selection [3, 4].

Acknowledgements

This research was supported by the European Union and co-financed by the European Regional Development Fund within the project TÁMOP-4.2.1/B-09/1/KONV-2010-0005.

References

[1] Katalin Kopasz, Péter Makra, Zoltán Gingl, A transparent, open-end and open-source measurement solution in natural science education, *Eur. J. Phys. 32*, pp. 491–504, March 2011.

[2] `http://www.noise.physx.u-szeged.hu/edudev/edaq530`

[3] Horst Bunke, János Csirik, Zoltán Gingl, Erika Griechisch, Online signature verification method based on the acceleration signals of handwriting samples, *Lecture Notes in Computer Science*,Springer 2011, Volume 7042/2011, pp. 499–506

[4] János Csirik, Zoltán Gingl, Erika Griechisch, The Effect of Training Data Selection and Sampling Time Intervals on Signature Verification, *First International Workshop on Automated Forensic Handwriting Analysis (AFHA 2011)*, Online Proceedings, pp. 6–10, `http://ceur-ws.org/Vol-768`

The vagueness measure: a new interpretation and an application to image thresholding

Gergely Gulyás and József Dombi

Here, we introduce a new interpretation of the vagueness measure [1] and an application for this approach. The vagueness measure is a generalized fuzziness measure based on a class of fuzzy operators. These operators contitute a coherent system which is called Pliant system [2], a subclass of continuous-valued logic. In this class each operator is defined by one generator function. The new property of the vagueness measure proposed in this work is that if the vagueness measure is computed for the distribution function of a given population, the value obtained gives a similar characteristic as the standard deviation of the population. Based on this feature, a new global thresholding algorithm was developed that generalizes the idea of Otsu's optimality criterion by the means of continuous-valued logic. The advantage of this approach is that it can be tuned for different problems by applying the appropriate operators. The performance of this method is compared with other commonly used algorithms to validate the usefulness of the proposed approach. The method was applied to a set of synthetic images for the sake of objectivism and we give a real application example where it is a part of a segmentation algorithm which segments skin lesions in digital dermoscopy images. Although the purpose of this algorithm is to threshold a grayscale image, it can be generalized for other tasks that require the separation of two or more populations, characterized by real values.

Acknowledgements

This study was partially supported by the TÁMOP-4.2.1/B-09/1/KONV-2010-0005 programme of the Hungarian National Development Agency.

References

[1] J. Dombi. Fuzziness Measure in the Pliant System: The Vagueness Measure, *Acta Technica Jaurinensis, Vol. 4, No. 1*, 2011.

[2] J. Dombi. Pliant operator system, *Recent Advances in Intelligent Engineering Systems, Studies in Computational Intelligence, Springer*, 2012.

Binary Tomography Using Two Projections and Morphological Skeleton

Norbert Hantos, Péter Balázs, Kálmán Palágyi

Binary Tomography (BT) [1] aims to reconstruct binary images from their projections. In the most common applications of BT usually just few projections of the object can be measured, since the acquisition of the projection data can be expensive or damage the object. Owing to the small number of projections the binary reconstruction can be extremely ambiguous. A common way to reduce the number of solutions of the reconstruction task is that certain geometrical properties (e.g., convexity and/or connectedness) are satisfied.

In this talk we investigate a new kind of prior information, the skeleton of the image to be reconstructed. Skeleton is a region-based shape descriptor which represents the general form of binary objects. One way of defining the skeleton of a 2-dimensional continuous object is as the set of the centers of all maximal inscribed (open) disks. The skeleton of a discrete binary image can be characterized via morphological operations [2], where disks are approximated by successive dilations of the selected structuring element that represents the unit disk. An interesting property of the morphological skeleton is that the original binary image can be exactly reconstructed from the skeletal subsets. In this work, we deal with the reconstruction problem in which the entire morphological skeleton (instead of the individual skeletal subsets) and two projections of the original image are known.

In the reconstruction process the prior knowledge is often incorporated into an energy function, thus the reconstruction task is equivalent to a function minimization problem. Many common methods exist to solve that kind of problems. In this talk, we show how to use Simulated Annealing (SA) [3] for the binary reconstruction problem using the information of the projections and the skeletal points. We show that, although theoretically the problem is non-unique, under some circumstances an acceptable image quality can be achieved. We propose three different methods to solve the above problem, based on parametric SA reconstruction.

This research was supported by the European Union and co-financed by the European Regional Development Fund within the project TÁMOP-4.2.1/B-09/1/KONV-2010-0005. The work of the second author was also supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences and by the OTKA PD100950 project of the National Scientic Research Fund.

References

[1] Herman, G.T., Kuba, A. (eds.): Advances in Discrete Tomography and Its Applications. Birkhäuser, Boston (2007)

[2] Gonzalez, R.C., Woods, R.E.: Digital Image Processing (3rd Edition). Prentice Hall (2008)

[3] Kirkpatrick, S., Gelatt Jr., C.D., Vecchi, M.P.: Optimization by Simulated Annealing. Science 220, 671–680 (1983)

On Some Subsystems of Interval-Valued Logic

László Hegedüs

The non-classical logic we consider here differs from classical propositional logic mainly in the interpretation function. A computing device, based on this idea can be quite powerful as shown in [3, 4, 5], also an application of this paradigm can be found in [6]. In classical propositional logic, an interpretation maps each propositional variable an element of the set $\{0, 1\}$. In the case of many-valued logics, variables can be assigned values from different sets (e.g. $\{0,\ 1/2,\ 1\}$, $\{1, \ldots, n\}$, or even $[0, 1]$). In the case of interval-valued logic (see for example: [1, 2]), a union of subintervals of $[0, 1]$ is assigned to each variable. We also require all of these subintervals to be in the form $[a, b)$. In other words an interval-value is a set of ordered pairs $(a_{i,1}, a_{i,2})$ $(i = 1, \ldots n)$ such that $0 \leq a_{i,1} < a_{i,2} < 0$, and $a_{i,1} \leq a_{j,1} < a_{i,2}$ is not possible for any distinct $i, j \in \{1, \ldots, n\}$. Logical operators negation, conjunction, disjunction and implication are defined as $\overline{A} = [0, 1) \setminus A$, $A \cap B$, $A \cup B$, $\overline{A} \cup B$ respectively, where A, B are interval-values. Some non-logical operators are also defined in [2], here we only mention $*$, the product operator that is

$$A * B = \bigcup \{[a_{i,1} + b_{j,1}(a_{i,2} - a_{i,1}), a_{i,1} + b_{j,2}(a_{i,2} - a_{i,1})) \mid 1 \leq i \leq n, 1 \leq j \leq m\},$$

where A and B consists of intervals $[a_{n,1}, a_{n,2})$ $(1 \leq n \leq k)$ and $[b_{m,1}, b_{m,2})$ $(1 \leq m \leq l)$ respectively. We note that this product is not commutative, but associative. An additional operator, based on Kleene's strict implication will be defined. One possible usage of this operator is to represent the semantic consequence relation.

We will investigate subsystems of interval-valued logics, where each value consists of only one interval. Thus the following two main cases are considered:

1. Each value starts from a designated end of $[0, 1)$. We have three sub-cases:
   All interval-values are in the form
   (a) $[0, a)$ where $0 < a \leq 1$,
   (b) $[a, 1)$ where $0 \leq a < 1$,
   (c) $[0, a)$ or $[b, 1)$ where $0 < a \leq 1$ and $0 \leq b < 1$.
2. Each value is in the form $[a, b)$, where $0 \leq a < b \leq 1$.

In each of these cases, some logical operators must be redefined in order to have a closed system. We will prove logical equivalences and show, that in all of these cases $\neg X \vee X$ is not necessarily a tautology. Thus some of them can be considered as intuitionistic logics.

References

[1] Nagy B.: *A general fuzzy logic using intervals*, 6th International Symposium of Hungarian Researchers on Computational Intelligence, Budapest, Hungary, 2005, pp. 613–624.

[2] Nagy, B.: *An Interval-valued Computing Device*, in CiE 2005: New Computational Paradigms, pp. 166–177.

[3] Nagy, B.: *Effective Computing by Interval-values*, in Proceedings of INES 2010, pp. 91–96.

[4] Nagy, B. and Vályi, S.: *Interval-valued computations and their connection with PSPACE*, in Theoretical Computer Science, **394**/3, 2008, pp. 208–222.

[5] Nagy, B. and Vályi, S.: *Prime factorization by interval-valued computing*, in Publicationes Mathematicae Debrecen, **79**/3-4, 2011, pp. 539–551.

[6] Zámbó, L. and Nagy, B.: *Optimization of the Painting Problem by a Genetic Approach using Interval-values*, CINTI 2011, pp. 127–132.

Periodicity of Circular Words

László Hegedüs and Benedek Nagy

We investigate some properties of circular words (or necklaces as mentioned in [6]). An ordinary word w is just a (finite or infinite) sequence of symbols (e.g. in the finite case: $w = w_1 w_2 \dots w_n$). A set of symbols are called an alphabet and usually denoted by Σ. For a detailed introduction to combinatorics on words the reader can consult [3].

A circular word is obtained from a finite word w by joining it at the two extremes (i.e. the beginning and the end of w). We denote it by w_o and it can also be viewed as the set of all conjugates of w (i.e., $w_o = \{v \in \Sigma^* \mid v \text{ is a conjugate of } w\}$). It can easily be seen that if w is a primitive word, then the number of conjugates of w is $|w|$. Results about pattern avoidance of circular words can be found in [1, 2, 5]. Also [4] contains some applications to integer sequences.

For traditional (i.e., linear) words, the definition of a period is as follows. Let $v \in \Sigma^*$ be a word and $r = \frac{p}{q} \in \mathbb{Q}$ such that $0 < r < 1$ and $p = |v|$. Then $v^r = v_1 v_2 \dots v_q$. A positive integer p is a *period* of w if there exist $v \in \Sigma^*$ such that $|v| = p \leq |w|$ and $w = v^r$ for some $r \in \mathbb{Q}$. We introduce new notions of *weak-* and *strong period*s of circular words and investigate their properties. The number $p > 0$ is a weak period of w_o if it is a period of some conjugate of w. Similarly, p is a strong period of w_o if it is a period of all conjugates of w_o.

A theorem of Fine and Wilf (see [3]) states that if p and q are both periods of a linear word w and $p + q - \gcd(p, q) \leq |w|$, then $\gcd(p, q)$ is also a period of w. This is not true for weak periods of circular words. Consider for example the word $(aabab)_o$ which has weak periods $2, 3, 4, 5$ and $2 + 3 - 1 \leq 5$, but $\gcd(2, 3) = 1$ is not a weak period of this word. Thus we investigate so called *paired period*s, that are in the form $(p, p + 1)$ such that p and $p + 1$ are both weak periods of w_o. Over the two letter alphabet, only words in the form $(ab)^*$ or $(ba)^*$ cannot have paired periods. The following list is a summary of some of our results:

- If w_o is a non empty circular word with strong period p, then p divides $|w|$. Corollary: If $|w|$ is a prime and there exists a strong period p of w_o such that $p < |w|$, then w is a unary word.
- Let $w \in \Sigma^*$ an arbitrary word. Then w_o has factors xx and yy with $|y| = |x| + 1$ iff $(|w| - |y|, |w| - |x|)$ is a paired period of w_o. A special case is when $xx = aa$ for some $a \in \Sigma$ and $yy = \varepsilon$ (the empty word). Then w has paired period $(|w| - 1, |w|)$.

Using our results we also state two conjectures about circular words obtained from the Thue-Morse and Fibonacci words:

- Circular words that are obtained from T_n (i.e., the nth Thue-Morse word) with $n \geq 5$ have only the paired periods that can be constructed from $\{\ell - 4, \ell - 3, \ell - 2, \ell - 1, \ell\}$, where $\ell = |T_n|$.
- Circular words that are obtained from F_n (i.e., the nth Fibonacci word) with $n \geq 4$ have only the paired periods that can be constructed from $\{k - 3, k - 2, k - 1, k\}$, where $k = |F_n|$.

References

[1] Currie, J. D. and Fitzpatrick, D. S.: *Circular words avoiding patterns.* in Proceedings of DLT 2002. pp. 319–325.

[2] Fitzpatrick, D. S.: *There are binary cube-free circular words of length n contained within the Thue-Morse word for all positive integers n* in Ars Combinatorica **74**, 2005.

[3] Lothaire, M.: Combinatorics on Words. Addison-Wesley, 1983.

[4] Rittaud, B. and Vivier, L.: *Circular words and applications*, in Proceedings of Words 2011, Electronic Proceedings in Theoretical Computer Science **63**, pp. 31–36.

[5] Shur, A. M.: *On ternary square-free circular words.* Electr. J. Comb. **17**, 2010.

[6] Smyth, W.: *Computing patterns in strings.* Addison-Wesley, 2003.

A Probabilistic Software Quality Model for C# – an Industrial Case Study

Péter Hegedűs

Both for software developers and managers it is crucial to have clues about different aspects of the quality of their systems. The information can mainly be used for making decisions, backing up intuition, estimating future costs and assessing risks. The ISO/IEC 9126 standard [1] defines six high-level product quality characteristics which are widely accepted both by industrial experts and academic researchers. These characteristics are: *functionality*, *reliability*, *usability*, *efficiency*, *maintainability* and *portability*. The characteristics are affected by low-level quality properties that can be *internal* (measured by looking inside the product, e.g. by analyzing the source code) or *external* (measured by execution of the product, e.g. by performing testing).

Maintainability is probably the most attractive, observed and evaluated quality characteristic of all. The importance of maintainability lies in its very obvious and direct connection with the costs of altering the behavior of the software. Although, the quality of source code unquestionably affects maintainability, the standard does not provide a consensual set of source code measures as internal quality properties. The standard also does not specify the way how the aggregation of quality attributes should be performed. These are not deficiencies of the standard, but it offers a kind of freedom to adapt the model to specific needs.

We have introduced a practical quality model in one of our previous works [2] that differs from the other models in many ways:

- It uses a large number of other systems as benchmark for the qualification.
- The approach takes into account different opinions of many experts, and the algorithm integrates the ambiguity originating from different points of view in a natural way.
- The method uses probabilistic distributions instead of average metric values, therefore providing a more meaningful result, not just a single number.

Although the introduced model proved to be useful and accepted by the scientific community, real industrial settings and evaluations are required to show that our solution is useful and applicable in real environments too. Additionally, the first published model is only a prototype for the Java language and weighted by a small number of researchers and practitioners. The goal of my presented work is to develop a method and model for estimating the maintainability of the C# systems of a large international company. To achieve this goal, the following tasks were completed:

- Together with the industrial partners, we have introduced a new maintainability model for systems written in the C# language.
- A benchmark from the C# systems of the company has been created (several millions of C# code lines has been analyzed).
- A method and tool has been developed for qualifying the smaller components of the company's software using the benchmark - producing a relative measure for maintainability of the components (we were able to rank the components of the company).
- A new weighting has been created involving the developers and managers.
- According to the method and model, a large number of components have been evaluated.

The results were discussed after the evaluation and compared to the developers' opinions. The industrial application of our method and model was successful as the opinions of the developers highly correlated with the maintainability values produced by our C# maintainability model. This result shows that our probabilistic quality model is applicable in the industry, as the industrial partners accepted the provided results and found our approach and tool very useful.

Acknowledgements

This research was supported by the Hungarian national grants GOP-1.1.1-11-2011-0038, GOP-1.1.1-11-2011-0006, and OTKA K-73688.

References

[1] ISO/IEC, ISO/IEC 9126. Software Engineering - Product quality. *ISO/IEC*, 2001.

[2] T. Bakota, P. Hegedűs, P. Körtvélyesi, R. Ferenc, and T. Gyimóthy. A probabilistic software quality model, *in proceedings of the 27th IEEE International Conference on Software Maintenance (ICSM)*, sept. 2011, p. 243-252.

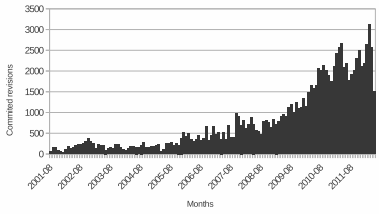
The Security History of the WebKit Browser Engine

Renáta Hodován

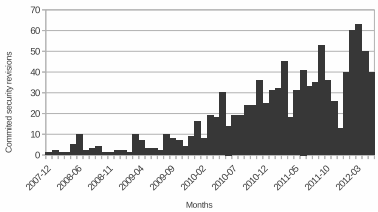
Over the years, more and more emphasis has been placed on the web domain. Customers are constantly requesting new features in web-applications and in the hosting browser itself as well. As these requests are fulfilled, the software becomes more and more complex and it's getting harder to take care of all side effects of the modifications. This can lead to undesired behaviour along with security vulnerabilities.

At the conference, I'll focus on the browser component and take a closer look at the security evolution of the WebKit layout engine. Among others, WebKit powers the Apple Safari and Google Chrome browsers and it holds nearly 36% of the browser market according to the survey of StatCounter in February 2012. Since WebKit is an open-source project, all implementation details and most bug entries are available. In this survey, I investigate a dataset that was generated from the publicly accessible security bug entries of WebKit. Furthermore, I determine a trend of the number of introduced security bugs over revisions and show the relation between the size of the code and the probability of the existing security holes.

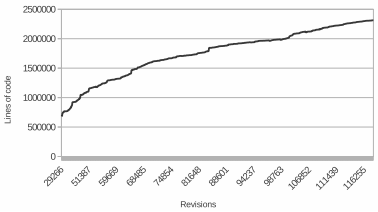
As Chart (a) shows, the number of committed revisions is growing increasingly over the months in the last few years. In line with this, the number of exposed security vulnerabilities follows the same trend, see Chart (b). This hangs together with the theorem that as a software is getting older and more complex [1], its maintenance and development turn into a hard challenge and we should count on undesired consequences. On Chart (c) we can see the way the size of the source code of the whole project is changing. And the last one shows how many lines of source code were modified during the months as a result of fixing security holes.



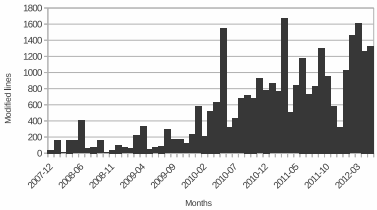
(a) Commited revisions over months



(b) Commited security bug fixes



(c) Code size over revisions



(d) The number of modified lines over security revisions

References

[1] Roger S. Pressman. *Software Engineering – A Practicioner's Approach*. McGraw-Hill, 5th edition, 2001.

Introduction to an automatized cheating prevention system at computerized open-book lab exams

Balázs Keresztury and László Cser

Organizing exams puts heavy load on the responsible departments at almost every university. To be fully prepared for fair and professional examinations, the department has to create tests, evaluate them by multiple teachers, print them, securely distribute them and after the actual examination they also have to keep all the documentation for many years. During the actual exams departments have to provide sufficient number of exam supervisors who keep an eye on the students while they are working on their tests. This is especially important when students write open-book exams, where they are allowed to use almost every help they can find online. Even if they can freely use online tutorials, forums or slides, they are strictly forbidden to utilize any kind of manual help, which typically includes instant messengers, email systems, etc. These, seemingly contradictory requirements forced us to research this area and come up with a viable solution.

Our department is responsible for organizing big amount of exams each year at Corvinus University of Budapest (CUB), which let us realize that many of the tasks above could be significantly simplified by using a Learning Management System (LMS). Since our university successfully adopted the widely used and accepted Moodle LMS[1, 3], we decided to use it to support our examination process. However, switching to computerized exams not only brought advantages, but also some disadvantages. While on one hand exams generally became safer (Moodle is capable of restricting users by authentication, quiz passwords, ip ranges, etc), on the other hand we faced the problem that users were able to save exam questions without any effort, which rendered them unusable later. Even if this problem could be fixed by some simple measurements, securing open-book exam was still a big challenge for us.

To face this challenge, our department decided to develop a software system which was responsible for making electronic open-book exams more secure. During the specification of requirements many ideas came up, including various known methods of cheating[4] and possible ways to prevent them[2]. After the usual specification of functional and non-functional requirements, we came to the following conclusion: The system has to be rock-solid, easy to use, multilingual, integrated with Moodle and it should be highly customizable. To achieve these goals, we have researched the market for similar solutions, chosen the programming and database platform, and started the actual development process. The implementation was followed by smaller test runs at optional pre-exams, but later we significantly broadened the circle of users. In the first runs we only typically added some features based on the feedback of the users, but later we moved the emphasis to stress-testing.

In this talk we plan to tell more about why one needs a system like this, if there is any similar solution available from the market, who benefits from using it, give better insight to our development process and share our experiences.

Acknowledgements

The financial support of TÁMOP 4.2.2/B-10/1-2010-0023 is gratefully acknowledged.

References

[1] Byoung-Chan Lee, Jeong-Ok Yoon, and In Lee. Learners' acceptance of e-learning in South Korea: Theories and results. *Computers & Education*, 53(4):1320–1329, 2009.

[2] E J SAUNDERS. CONFRONTING ACADEMIC DISHONESTY. *JOURNAL OF SOCIAL WORK EDUCATION*, 29(2):224–231, 1993.

[3] Erik M van Raaij and Jeroen J L Schepers. The acceptance and use of a virtual learning environment in China. *Computers & Education*, 50(3):838–852, 2008.

[4] Kevin M Williams, Craig Nathanson, and Delroy L Paulhus. Identifying and Profiling Scholastic Cheaters: Their Personality, Cognitive Ability, and Motivation. *JOURNAL OF EXPERIMENTAL PSYCHOLOGY-APPLIED*, 16(3):293–307, 2010.

Wavelets and rational function methods in ECG signal processing

Péter Kovács

The electrocardiograms (ECG) represent the electrical activity of the heart related to different phases of the muscular contraction. These electrical changes can be measured on the surface of the skin by electrodes. An ECG signal records the potential difference between these electrodes. The electrocardiogram consists of five standard waves labelled with the letters P, Q, R, S, T. The amplitudes, durations and shapes of these waves carry important diagnostic information about the patients. There is a wide range of recent applications of digital signal processing including detection, analyzation and compression.

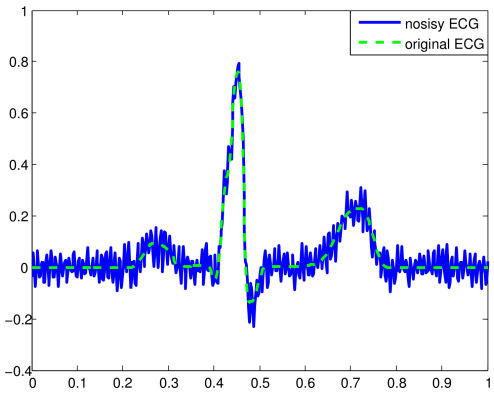
Lately, we used rational functions to model the electrocardiogram [3]. In order to outline this method, let us take the elementary rational functions on the unit disc $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$

$$r_a(z) := \frac{1}{1 - \overline{a}z} \qquad (|z| < 1)\ ,$$

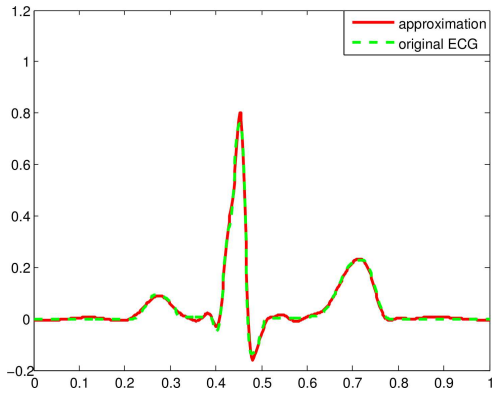
where the pole is $a/|a|^2$. a itself is called the inverse pole of r_a . Then we take linear combinations of powers of elementary functions and consider them on the unit circle. We use these functions to model the ECG signal f as follows

$$f \approx \sum_{n=1}^{N} \sum_{j=1}^{n_k} c_{n,j} \cdot r_{a_n}^j \qquad (N, n_k \in \mathbb{N}),$$

where the $c_{n,j}$'s are complex coefficients. Three different poles turned to be appropriate in most cases [2]. In the present work we compare the widely used wavelet-based techniques [1] with those of rational functions [3]. We compared these methods in various aspects.



(e) Synthetic ECG signal with an 50Hz power line interference.



(f) Denoised ECG curve smoothed by basic rational functions.

Figure 2: Rational function approximation of a synthesized electrocardiogram.

Firstly, we compare the rate of compression by rational functions in contrast with the Daubechies and Morlet wavelets. Furthermore, we want to use an objective measure of level of compaction and distortion rate. For this reason, synthesized ECG signals [5] were used to test these methods. In this case, we can generate electrocardiograms with different diagnostical and geometrical features, but we can get the original signals in analytic forms as well. So, both compression and distortion rate can be measured by using different types of error measures, such as percentage root mean square difference (PRD) and weighted diagnostic distortion (WDD) [6]. By taking advantage of the analytic form of the original signal, one can test the sensitivity of these methods by changing the signal-to-noise ratio (SNR). On Figure 2(f) one can see the

reconstructed signal marked by red line. The original ECG curve has 512 sample, but we used only 16 complex coefficients and 3 complex poles for the representation.

Secondly, a related problem is the filtering property. Namely, these methods can be used for denoising ECG signals. Different types of noise such as muscle noise, baseline drift, etc. were used to test the smoothing ability of rational functions in contrast with Daubechies and Morlet wavelets. During filtering it is important to keep the most significant ECG features by tempering the distortion rate of the process. An example for rational filtering is given in Figure 2.

Finally, we propose a wave detection algorithm based on the rational function approximations. This method separates one period of an electrocardiogram into three main parts that are designated alphabetically as P, QRS and T waves. On one hand, it is a rough approximation of the original signal, because we use only the most significant terms from the projection (2). On the other hand, the onsets and the offsets of the main lobes of the ECG curve can be predicted well. Furthermore, both synthesized electrocardiograms and real ECG databases such as physionet [4] were used to test and compare the accuracy of this wave detection algorithm with wavelet-based methods. These databases are also annotated, so the test results can be easily achieved.

Acknowledgements

The Project is supported by the European Union and co-financed by the European Social Fund (grant agreement no. TÁMOP 4.2.2./B-10/1-2010-0030).

References

[1] P. S. Addison, Wavelet transforms and ECG: a review, *Physiological Measurement,* vol. 26, pp. 155–199, 2005.

[2] S. Fridli, P. Kovács, L. Lócsi and F. Schipp, Rational modeling of multi-lead QRS complexes in ECG signals, *Annales Univ. Sci. Budapest., Sect. Comp.*. vol. 37, pp. 145–155, 2012.

[3] S. Fridli, L. Lócsi and F. Schipp, Rational function system in ECG processing, *Proc. 13th EUROCAST 2011, Part I (eds. R. Moreno-Díaz et al.)*, LNCS 6927, pp. 88–90, 2011

[4] A.L. Goldberger, L.A.N. Amaral, L. Glass, J.M. Hausdorff, P.Ch. Ivanov, R.G. Mark, J.E. Mietus, G.B. Moody, C.K. Peng and H.E. Stanley, PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. Circulation 101(23):e215-e220, 2000. `http://circ.ahajournals.org/cgi/content/full/101/23/e215`

[5] P. Kovács, ECG signal generator based on geometrical features, *Annales Univ. Sci. Budapest., Sect. Comp.*. vol. 37, pp. 247–260, 2012.

[6] Y. Zigel, A. Cohen and A. Katz, The Weighted Diagnostic Distortion (WDD) Measure for ECG Signal Compression, *IEEE Transactions on Biomedical Engineering*, vol. 47, pp. 1422–1430, 2000.

PhP + MySQL questions and possible answers

Edit Lázár

During the education of programming we meet varios problems. In my essay I want to show some problems and their possible solutions in PhP and MySQL. Today these two programs together make a very fast and useful complex on web-programming. I want to pay attention to the possible errors and failures what can programmers do writing a questionnary-make program. I will make a process from the planning of software and database to the realize. I will give a list about problems, their possible solutions, if I know they, and the source where I found the solution (e.g. character-set using problems, random-procedure, etc.). I won't be able to enumarate all of problems and solutions but I intended to show a lot and I want to give some help to teachers, too.

Hybrid Face Detector Based on Boosted Classifiers

Szidónia Lefkovits

The task of object detection is one of the most widespread research domains in computer vision. Despite of the fact that humans have the extraordinary capacity to detect several types of objects or object categories, the general problem of object detection has not been solved yet. Nowadays, a high-performance system supposes a high rate of detection, and at the same time, a low rate of false detections.

This paper presents a robust face detection system that combines holistic and component-based methods. At first, a very fast appearance-based method detects the face as a whole using the well-known AdaBoost algorithm [1, 2]. This method does an exhaustive search over the image in different scales and positions using the image pyramid. Millions of image parts have to be analysed, out of which only a few are faces and all the others are backgrounds. A rate of $5 \cdot 10^{-6}$ false detections is not sufficient to eliminate all of the backgrounds [3]. Visually, the falsely detected images are not similar at all to the face. This is the reason why in the second phase, we propose a detection system whose role is to examine the image in detail. This process is slow, but it needs to be done only rarely.

The second part of the system is built with a well-known deformable object model [4] based on our proposed local descriptor. The deformable model solution can be implemented through a dynamic programming technique [5] which applies a new proposed algorithm for general distance transform [6]. The most important factor that determines the system's performance is the local image descriptor. Our idea comes from the working principle of the visual receptive field of mammalians. Hence, the proposed descriptor is based on Gabor filters. As you may know, the bi-dimensional Gabor filter [7] is defined in a high dimensional space of 9 parameters. One of the difficult tasks is to limit the domain of parameters. The first step is a thorough study that defines their theoretical limits. The fine-tuning process of the parameters for one kind of patch is done experimentally, with the goal to retrieve the most characteristic set of filters for the target object. The selection of the filters is done by the GentleBoost algorithm [8]. We implemented a supervised learning process that needs a huge amount of image patches of the object. This translates to a high computational complexity of the proposed descriptor. Nevertheless, it can be implemented in almost all detection systems.

The case study has been made for face detection, and the detailed search has used the above-mentioned descriptor for eye detection. We demonstrate the viability of the proposed system experimentally. We try to find the optimum between local and global aspect-based methods and between computational speed and detection performances.

References

[1] M. J. P. Viola. Fast multi-view face detection. Technical Report TR2003-096, *Mitsubishi Electric Research Laboratories, Cambridge*, 2003

[2] R. Lienhart, L. Liang, and E. Kuranov. A detector tree of boosted classifiers for real-time object detection and tracking. *In IEEE ICME2003*, pages 277-280, 2003

[3] Sz. Lefkovits. Assessments of building classifiers for face detection. *In International Conference On Recent Achievements in Mechatronics, Automation, Computer Sciences and Robotics MACRO*, pages 175-186, 2009.

[4] M. A. Fischler and R. A. Elschlager. The representation and matching of pictorial structures. *In IEEE Transactions on Computers*, volume C-22, pages 67-92, 1973

[5] P. F. Felzenszwalb and D. P. Huttenlocher. Pictorial structures for object recognition. *International Journal of Computer Vision*, 61(1): 55-97 ,2005

[6] Sz. Lefkovits. Numerical computation method of the general distance transform. *In KEPT 2011 International Conference Knowledge Engineering Principles and Techniques*, pages 143-152, 2011

[7] J. R. Movellan. Tutorial on Gabor filters

[8] J. Friedman, T. Hastie, and R. Tibshirani. Additive Logistic Regression: a Statistical View of Boosting. *The Annals of Statistics*, 38(2), 2000

Verified localization of trajectories with prescribed behaviour in the forced damped pendulum

Balázs László Lévai

In mathematics, it is quite difficult to define exactly what chaos really means. In particular, it is easier to prepare a list of properties which describe a so called chaotic system than give a precise definition. A dynamic system is generally classified as chaotic if it is sensitive to its initial conditions. Chaos can be also characterized by dense periodic orbits and topological transitivity.

While studying computational approximations of solutions of differential equations, it is an important question is whether the given equation has chaotic solutions. The nature of chaos implies that the numerical simulation must be carried out carefully, considering fitting measures against possible distraction due to accumulated rounding errors. Unfortunately except a few cases, the recognition of chaos has remained a hard task that is usually handled by theoretical means [3].

In our present studies, we investigate a simple mechanical system, Hubbard's sinusoidally forced damped pendulum [3]. Applying rigorous computations, his 1999 conjecture on the existence of chaos was proved in Bánhelyi *et al.* [1] in 2008 but the problem of finding chaotic trajectories remained entirely open. This time, we present a fitting verified numerical technique capable to locate finite trajectory segments theoretically with arbitrary prescribed qualitative behaviour and thus shadowing different types of chaotic trajectories with large precision. For example, we can achieve that our pendulum goes through any specified finite sequence of gyrations by choosing the initial conditions correctly.

To be able to provide solutions with mathematical precision, the computation of trajectories has to be executed rigorously. Keeping in mind this intention, we calculated the inclusion of a solution of the differential equation with the VNODE algorithm [5] and based on the PROFIL/BIAS interval environment [4]. The search for a solution point is a global optimization problem to which we applied the C version of the GLOBAL algorithm, a clustering stochastic global optimization technique [2]. This method is capable to find the global optimizer points of moderate dimensional global optimization problems, when the relative size of the region of attraction of the global minimizer points are not very small.

References

[1] BÁNHELYI, B., T. CSENDES, B.M. GARAY, AND L. HATVANI, *A computer–assisted proof for Σ_3–chaos in the forced damped pendulum equation*, SIAM J. Appl. Dyn. Syst., 7, 843–867 (2008).

[2] CSENDES, T., *Nonlinear parameter estimation by global optimization – efficiency and reliability*, Acta Cybernetica, 8, 361-370 (1988).

[3] HUBBARD, J.H., *The forced damped pendulum: chaos, complication and control*, Amer. Math. Monthly, 8, 741–758 (1999).

[4] KNÜPPEL, O., *PROFIL – Programmer's Runtime Optimized Fast Interval Library*, Bericht 93.4., Technische Universität Hamburg-Harburg (1993).

[5] NEDIALKOV, N.S., *VNODE – A validated solver for initial value problems for ordinary differential equations*, Available at `www.cas.mcmaster.ca/~nedialk/Software/VNODE/VNODE.shtml` (2001).

Analysis of a hyperbolic optimization method

Levente Lócsi, Ferenc Schipp

In our works concerning the analytic representation of ECG signals using complex rational functions we had to face an optimization problem which is defined on the complex unit disc, i.e. on the set $\{ z \in \mathbb{C} : |z| < 1 \}$. It turned out that the function to be optimized has several local minima, sophisticated methods are required to reach an optimal solution. We have developed and analysed many algorithms suitable for this problem, see e.g. [1] and [2].

This work to be presented here is aimed at the study of another algorithm, which is based on the hyperbolic version and variation of the well-known gradient method. Note that the complex unit disc also provides the so-called Poincaré disc model of the Bolyai–Lobachevsky hyperbolic geometry. We will also make advantage of the fact that the congruent transforms in this geometry can be expressed by the means of some rational complex functions, namely the Blaschke functions, defined as

$$B_{\delta,a}(z) := \delta \frac{z-a}{1-\overline{a}z}. \qquad (\delta, a, z \in \mathbb{C}, |\delta| = 1, |a| < 1, z \neq 1/\overline{a})$$

Our aim is to create a MATLAB implementation of this proposed algorithm, to present some numerical experiments and to analyse the method and its convergence properties in the case of some simple test functions.

The Project is supported by the European Union and co-financed by the European Social Fund (grant agreement no. TAMOP 4.2.1./B-09/1/KMR-2010-0003).

References

[1] S. Fridli, L. Lócsi, F. Schipp, Rational Function Systems in ECG Processing, *Proc. 13th Int. Conf. Computer Aided Systems Theory (EUROCAST 2011), Part I (eds. R. Moreno–Díaz et. al.),* Springer LNCS 6927 (2011), 88–95.

[2] S. Fridli, P. Kovács, L. Lócsi, F. Schipp, Rational Modeling of Multi-lead QRS Complexes in ECG Signals, *Annales Univ. Sci. Budapest., Sect. Comp.,* **37** (2012), 145–155.

Adapting the Reactive Search Optimization and Visualization Algorithms for Multiobjective Optimization Problems; Application to Geometry

Amir Mosavi, Miklos Hoffmann, A. S. Milani

In most of the real-world optimal design problems of engineering and business processes, in order to improve the functionality, the operating parameters need to be accurately tuned with the aid of the multiobjective optimization algorithms for which many conflicting objectives have to be traded off in selecting the preferred solution(s). For solving the complicated multiobjective optimization problems numerous biology-inspired metaphors e.g. evolutionary algorithms [1] which have indeed a very limited learning capabilities, have been widely utilized so far.

On the other hand very recently the effectiveness of reactive search optimization (RSO) algorithms [3, 9] along with the visualization tools [2], in operations research and mathematical programming, covering a variety of different applications to multiobjective optimization, are becoming increasingly popular. The RSO algorithms, considered as the most advanced Brain-Computer Optimization (BCO) algorithms, are developed on the basis of involving the decision maker interactively in the loop; loading the intelligent expertise to the algorithm leading to increasing the learning capabilities. RSO employs learning for optimization, via integration of sub-symbolic machine learning techniques into the search heuristics so that the algorithm selection, adaptation and integration, are done in a rather automated way. In this way the past history of the search could be used for self-adaptation in an automated manner. Moreover visualization and optimization are connected through decision maker interaction.

In this article the RSO algorithms are compared with the evolutionary algorithms in terms of the effectiveness, usability, and efficiency. To do so three of the well-known geometrical optimization problems, previously solved by evolutionary algorithms, i.e. multiobjective optimization the curves and surfaces [4, 10], skinning problem [5], and welded beam design [6], are reconsidered utilizing RSO and visualization tools via a commercial implementation [7, 8]. The results are further carefully discussed in each study case.

References

[1] C. A. Coello Coello, G. Toscano, A micro-genetic algorithm for multiobjective optimization, in Lecture Notes in Computer Science no. 1993, E. Zitzler, K. Deb, L. Thiele, C. A. Coello Coello, and D. Corne, Eds. Berlin, Germany: Springer-Verlag, 2001, Proc. 1st Int. Conf. Evolutionary Multi-Criterion Optimization, pp. 126–140.

[2] C.V. Jones. Feature Article–Visualization and Optimization. INFORMS Journal on Computing, 6(3):221, 1994.

[3] R. Battiti, M. Brunato, F. Mascia, Reactive Search and Intelligent Optimization. Operations research/Computer Science Interfaces. Springer Verlag, 2008.

[4] R. Goldenthal, M. Bercovier, Design of Curves and Surfaces by Multiobjective Optimization, Mathematical Methods for Curves and Surfaces, 2004.

[5] R. Kunkli, M. Hoffmann, Skinning of circles and spheres, Computer Aided Geometric Design, 27, 611-621, 2010.

[6] G.V. Rekliatis, A. Ravindrab, K.M. Ragsdell, Engineering Optimisation Methods and Applications. 1983, New York: Wiley. M. Brunato, R. Battiti, Grapheur: A Software Architecture for Reactive and Interactive Optimization, in Lecture Notes in Computer Science, no. 6073, C. Blum and R. Battiti, Eds. Berlin, Germany: Springer-Verlag, 2010, LION 4, LNCS, pp. 232–246.

[7] A. Mosavi(2010). Multiple criteria decision-making preprocessing using data mining tools. IJCSI International Journal of Computer Science Issues, V7, Issue 2, No 1.

[8] A.Mosavi, M.Azodinia, Kasun N. Hewage, Abbas S. Milani, M.Yeheyis, Reconsidering the Multiple Criteria Decision Making Problems of Construction Workers Using; Grapheur, Poster presented at EnginSoft International Conference, Verona, Italy 20-21 Oct.

[9] A. Mosavi, Multiobjective Optimization of Spline Curves using modeFRONTIER, Proceedings of international conference on engineering optimization and International modeFRONTIERUsers– Meeting, Trieste, Italy, 2010.

[10] A. Mosavi, An Interactive Multiobjective Optimization and Decision Making Software Architecture with Engineering Applications,IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 4, July 2012.

[11] A.Mosavi, The large scale system of multiple criteria decision making; pre-processing, Large Scale Complex Systems Theory and Applications, 2010, Volume 9, Part 1.

Approximate dictionary matching for biomedical information extraction

György Móra

Information extraction techniques are widely used in various text mining fields like scientific literature mining and Web mining. Named entity recognizers often make use of word lists, that they subsequently map to the text. There are efficient methods available for mapping large dictionaries to texts. The species name recognizer Linnaeus, for instance, uses regular-expression-like patterns generated from a simple list of organism names [3]. The patterns are generated in such a way that the system can recognize spelling variants, but their flexibility is limited. Biomedical entities display a great spelling variability, hence the tagging of gene and protein names is essential for good information extraction in the biomedical domain. We developed a method for effectively matching large dictionaries containing millions of entity names using a flexible matching procedure. Our system was developed for gene and protein name matching, but it can also be used for other tasks where flexible matching is required.

Our approach tries to overcome the difficulties of biomedical entity matching. The order of the words in these entities may be different and extra punctuation characters and words may appear inside entity names. For instance, *member 3 of the solute carrier family 6* can be written as *solute carrier family 6, member 3*. Numbers can be written as Roman or Arabic numerals, which increases the number of possible variations. Greek letters are often transcribed or a similar Latin letter is used. For instance, *nuclear factor kappa beta* can be written as *NF-kappa B* or *NF-*$\kappa\beta$. To match all linguistic forms of a word we determined the base form of all words used in all dictionary entries and during the matching the base forms of the words in the texts are matched against the base form of the entries in the dictionary. Biomedical entities often have nonconventional plural or adjective forms. In our experiments, we used the LVGTools [1] linguistic package to determine the possible base forms of the words in the dictionary and in the text. Most words have one possible base form (like *dogs -> dog*), but there are ambiguous cases where only a syntactic or semantic analysis can determine which base form belongs to the word form in a given context. The LVGTools also handle the above-mentioned issues with Greek letters and Roman numerals as it generates all possible 'interpretations' of the ambiguous word.

Our approach utilizes the Lucene search engine [2] for indexing. The dictionary entries are split into words and the possible base forms of the words are determined. The entities are indexed with respect to the base forms and the number of words in the entity name. When the text is tagged, the possible base forms of the words are also determined. If one of the base forms of a word is contained in the index, the word is marked as a potential part of an entity name. Based on these annotations, a value for each word in each sentence is calculated which tells us how many words the longest entity name can have, which the current word may be a part of. This value is used as a maximum value when the possible names are queried from the index. This restriction speeds up the process because an entity name may be as long as 22 words. The tagging speed is approximately one document per second on biomedical full articles with over 7 million different entity names using a computer with a 2 GHz processor. The spans of the hits may overlap and a rule-based system filters out the matches to produce the longest non-overlapping annotations.

We developed a dictionary lookup tagger, which can be effectively applied in biomedical information extraction pipelines, to perform string matching with the necessary flexibility. We showed that this task can be completed in reasonable time with alternative indexing and searching methods.

References

[1] *SPECIALIST Lexicon and Lexical Tools*. National Library of Medicine (US), September 2009.

[2] Cutting, D. Lucene, 2001.

[3] Gerner, Martin, Nenadic, Goran, and Bergman, Casey M. LINNAEUS: a species name identification system for biomedical literature. *BMC bioinformatics*, 11(1):85, January 2010.

Wikipedia-based methods to identify noun compounds in running texts

István Nagy T.

In natural language processing, multiword expressions (MWEs) have been received special interest. Noun compounds (NCs) form a subtype of multiword expressions: they form one unit the parts of which are meaningful units on their own, and the unit usually has some extra meaning component compared with the meanings of the original parts [2]. The semantic relation between the parts of the noun compound may vary: it may express a "made of" relation (*apple juice*), a "location" relation (*neck pain*), a "made for" relation (*hand cream*) just to name a few. Thus, noun compounds encode some important meaning relations that can be fruitfully applied by e.g. information extraction systems. However, such applications require that noun compounds should be previously known to the system.

Noun compounds are very frequent in language use (in the Wiki50 corpus [2] 67.3% of the sentences contain a noun compound on average). Furthermore, they are productive: new noun compounds can enter the language all the time hence they cannot be exhaustively listed and appropriate methods should be implemented for their identification. In this paper, we introduce several methods to automatically identify noun compounds on the basis of Wikipedia, like dictionary labeling, rule based methods and machine learning based approaches and show how the expansion of Wikipedia helps the performance of different NC-detecting methods.

References

[1] Sag, I.A., Baldwin, T., Bond, F., Copestake, A., Flickinger, D.: Multiword Expres- sions: A Pain in the Neck for NLP. In: Proceedings of the 3rd International Confer- ence on Intelligent Text Processing and Computational Linguistics (CICLing-2002, Mexico City, Mexico (2002) 1-15.

[2] Vincze, V., Nagy T., I., Berend, G.: Multiword expressions and named entities in the Wiki50 corpus. In: Proceedings of RANLP 2011, Hissar, Bulgaria (2011).

Online Signature Feature Extraction from Video

József Németh

Handwritten signature verification is an intensively investigated area through it is an easy way to electronically verify a person's identity. While other authentication techniques may be more reliable, HSV is more convenient and less intrusive. Most early works focused only on the offline (static) HSV which only requires only an image of the signature. When the signatures are acquired using an electronic graphic tablet or an other special device the dinamics (i.e. the movement) of the writing is also captured and can be used for the verification (online systems, see [1]).

We present a video based online signature verification method. The signatures are acquired using a low-cost camera which records not only the track of the pen tip but other data like the movement of the whole pen. We use calibration patterns to estimate the camera parameters thus the trajectory of the pen can be reconstructed. Furthermore we extract other features which are also used to compare the signatures data.

We performed our experiments on our private database in which each person provided 10 genuine and 5 forgery signatures. The learning dataset is consisted of 5 randomly selected genuine signature for each signer. The comparison and classification of the signatures are achived using classical online HSV methods like dynamic time warping (DTW). These experiments showed that which features can be used most efficiently for signature verification. The proposed system achieved an equal error rate of 6%.

References

[1] G.K. Gupta. The State of the Art in On-line Handwritten Signature Verification, 2006.

Analysing Synchronization Contract Support in Modern Software Development Methodologies

György Orbán

The quality of the software systems must be increased. There are many techniques and methods to achieve this. In our paper we will focus one of these techniques the contract based development. Bertrand Meyer devised the expression "design by contract" which is a trademark of Eiffel Software. He created the Eiffel programming language in mind of contract support in the implementation and design. Not just a programming language, a software development methodology was created.

Four levels (syntactic, behaviour, synchronization, quality of service) of contracts can be separated based on Beugnard's categorization. At the different levels different type of contracts can be defined based on the specification. Contract based development is a continuously evolving area, but most of the developments are connected to the second, behaviour level. There are many tools which support contract based development at the implementation level. Programming languages like Eiffel, D with native support and other languages (Java, .NET, Python etc.) with extensions. There are some tools and extensions which can be integrated into an IDE and others can be used as a library. With these extensions the contracts can be weaved with different techniques into the source code or byte code. Contracts can be created at different levels in a software system. In the modern software development methodologies like Model Driven Development (MDD) the contract creation support should be at the modelling level. To achieve this we have different tools and methods. One of the most widely used modelling language is the Unified Modeling Language (UML). Object Constraint Language (OCL) expressions in the UML models can be used as contracts. With tools like DresdenOCL it is possible to generate Java source code from the OCL expressions. With this source code generation second level (behaviour) contracts can be created. But this is only one level from the four. There are many Interface Description Languages (IDL) for different types of systems and in most cases the source code generation is possible from these descriptions. The first two contract levels have more modelling and tool support.

In the paper we would like to focus on the modelling and implementation support for the synchronization (third level) contracts. From the modelling view the OCL, as a synchronization level contract definition language and from the implementation view different concurrency techniques will be examined. As a summary of our work we would like to recommend a software development process workflow with tools and methods used in the different steps to create more reliable, better quality software systems.

Acknowledgements

This research was supported by the European Union and co-financed by the European Social Fund (grant agreement no. TAMOP 4.2.1./B-09/1/KMR-2010-0003).

References

[1] Meyer, B., Object-Oriented Software Construction, Second edition, *Prentice Hall.* 1997

[2] Beugnard, A. et al., Making Components Contract Aware. *Computer*, 32(7), pp.38-45. 1999

[3] Jézéquel, J.-M., Object-oriented software engineering with Eiffel, Redwood City, CA, *Addison Wesley Longman Publishing Co., Inc.* 1996

[4] Morandi, B., Bauer, S.S. & Meyer, B., SCOOP - A Contract-Based Concurrent Object-Oriented Programming Model. *In P. Müller, ed. Advanced Lectures on Software Engineering. Berlin, Heidelberg: Springer Berlin Heidelberg,* pp. 41-90. 2010

[5] Sendall S., Strohmeier A., Specifying Concurrent System Behavior and Timing Constraints Using OCL and UML. *In Proceedings of the 4th International Conference on The Unified Modeling Language, Modeling Languages, Concepts, and Tools,* Springer-Verlag, London, UK, 391-405. 2001

Change management applying comparison of different versions of Dynamics AX

István Orosz and Tamás Orosz

Microsoft Dynamics AX continuously extends functional and technical approaches. Former Axapta and Navision applications are being replaced with the modern and smart AX and NAV solutions, like the other Dynamics modules and components. A brand new MS Dynamics implementation follows standard methodologies of Sure Step. However, companies, where one can find former Microsoft ERP systems, like Axapta or Navision, face to several change management issues and problems. One can figure out that business requirements for Software change management depend on specific business functions, industry, technologies and so on.

An organizational change management usually comes with organizational developments as well. The purpose of change management is to minimize the negative impacts on the organization and to avoid concerns. There are some different types of change management: changing the behavior of the personnel; technological changes; operational and structural changes; strategic and mission changes. The changes, which affect Dynamics AX, come from operational, structural and technological sides. The organizational change management should start with detailed description of the current situation, and after that focusing on the need of change and the ability of change. The objectives, content, and the process all should be part of the organizational change management plan. It uses the following techniques: performance metrics such as financial results, operational efficiency, leadership commitment, communication effectiveness. The perceived need for change to plan the appropriate strategies, in order to avoid change failures or solve troubled change projects. It is also important to make a clear governance and organizational model of the company, which is suited and aligned to the expectations of the company

This paper tries to introduces an agile way for business change management in Microsoft Dynamics environment and determine its usability in changing business environment. The comparison of different business requirements is made in different MS Dynamics business change issues.

Sustainability is one of the core factors of developing Dynamics AX 2012. One of the key aspects was to speed up the business and technological process of upgrading from the previous versions to AX 2012. Sure Step is the official software process for Dynamics AX. This is the tool of defining process phases, milestone roles, artifacts, cross phase processes, additional project management processes. Agile implementation model is an iterative, incremental process for developing Microsoft Dynamics Solutions. This Project Type gives customers greater control over the final solution because they can quickly change the direction of solution development and implementation from one sprint cycle to the next. It focuses delivering the whole functionality in smaller chunks, within a series of smaller developing series. By delivering and accepting smaller parts of development, the risk of the customer will not get the solution needed is smaller.

Our study tries to discover the options to build up the change management process flows within AX systems, and check the possible ways of creating them. We try to take a step forward to extend these workflows by using test driven software development methods.

Keywords: Change Management, Microsoft Dynamics AX, CRM, GP, NAV.

Empirical Studies of Reconstructing hv-Convex Binary Matrices from Horizontal and Vertical Projections

Zoltán Ozsvár and Péter Balázs

Tomography is a method of producing a three-dimensional image of the internal structure of an object from its projections, without damaging it. In $binary\ tomography$ we assume, that the examined object is homogeneous. In order to reduce the number of projections needed to the reconstruction, we further assume that the image satisfies certain geometrical conditions, such as hv-convexity.

The reconstruction of hv-convex binary matrices from their horizontal and vertical projections is proved to be NP-hard. In this paper, we take a closer look at the difficulty of the problem. We investigate different heuristic reconstruction algorithms of the class, and study them from the viewpoint of runnung-time and reconstruction quality. In the experiments we use a large set of test data, with different size and number of components. We observe that for each studied algorithm the dissimilarity of the reconstucted and the original images depends on the number of the components, rather than the size of the image. Futhermore, the reconstruction time of the core-shell algorithm depends both on the size of the image and the number of its components, while the speed of the simulated annealing reconstruction is mostly determined by the number of so-called switching components present in the image.

Acknowledgements

The work of the second author was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences and by the OTKA PD100950 project of the Hungarian Scientific Research Fund.

Synergy between mobile terminal and cloud computing.

Krisztián Pándi and Hassan Charaf Dr.

The role of cloud computing, as a new area of computing, is growing intensively. Mobil e computer technology has also greatly evolved in the recent years. The application variety that can b e used in mobile terminals is emerging, putting strong requirements both on software and hardware side of the mobile terminal. In cloud computing application services are provided through the Internet. Distributed and cloud computing is recognized to b e a new area for solving performance issues. Mobile terminal can take advantage from cloud computing. To cope with these new resources and fulfill new quality and performance requirements a more sophisticated architecture and resource management is necessary. The basis of effective resource management is a precise knowledge of the hardware and software capabilities; currently available performance metrics must be adapted for mobile terminals. Performance and resource utilization of the mobile terminal must be identified. Performance metrics serve as an input for resource management, deciding which application where to run, on the mobile terminal, or in the cloud.

Mobile terminal can use cloud for solving performance issues, and to obtain richer user experience. The aim of the study is to present architecture of mobile resource management, that can utilize benefits of cloud computing. Performance measurement and usable metrics are mandatory for our later research: decision making mechanism implementation. The goal of the mechanism is to decide where the optimal place to run for a certain service/application is. On the mobile terminal itself, or on other nearby mobile terminal, or on public cloud computing server. Hence a performance and usage of the mobile terminal should be determined.

Emerging complexity of the application used in mobile terminals implicate harnessing of these extra performance resources. Applications with distributed components differ from traditional non distributed applications in numerous attributes, such as communication type and overhead, latency, concurrency etc. The task of proposed mobile terminal resource and service management is to decide which application and service where to execute. To effectively fulfill this complex task; sophisticated and dedicated decision formula is needed. As such a formula amplifies adapted and dedicated software and performance metrics. Mobile terminal coupled with distributed system can be dynamic, changing over time, resulting CPU and network load changing. Therefore mobile terminal as a part of the distributed hierarchy needs to have metrics very different than traditional software and performance metrics [1].

With mobile computer technology progress, the software and hardware platform becomes more and more complex, together with the amount of the tasks meant to be processed. Mobile terminal has some special features in comparison with traditional computing; small size, dependence on limited battery lifetime, computing power is changing, possible presence of 3D hardware, network bandwidth is limited [2], and almost exclusively wireless, relatively small display size, the user input is special one.

Usually same or similar applications are used in mobile terminals as in traditional computers so similar user experience is expected. Therefore, with comparably less performance nearly the same look and feel is required. Thus, the capabilities of the mobile hardware should be efficiently harnessed with smart resource management and load balancing.

References

[1] Krisztián Pándi, PERFORMANCE MEASUREMENT ARCHITECTURE FOR MOBILE SOFTWARE SYSTEMS., In: XXVI. microCAD International Scientific Conference. Miskolc, Magyarország, 2012.04.29-2012.04.30. (Miskolci Egyetem) Miskolc: pp. 1-6. Paper I6. (ISBN: 978-963-661-773-8)

[2] Krisztián Pándi, Network Aggregation in Mobile Environment.,In: István Vajk, Renáta Iváncsy (szerk.),Proc. of Automation and Applied Computer Science Workshop (AACS). Budapest, Magyarország, 2010.06.25. Budapest: pp. 201-209.(ISBN: 978-963-313-004-9)

x86 instruction reordering for code compression

Zsombor Paróczi

Runtime executable code compression is a special method, which uses standard data compression methods and binary machine code transformations to achieve smaller file size, yet maintaining the ability to execute the compressed file as a regular executable. In a typical case the source code gets compiled with a compiler and linker, the output is a binary executable, which contains the machine code and the data needed by the application. This work focuses only on executable machine code for the Intel x86 instruction set.

Many compression methods for x86 machine code have been developed, most of them use model based compression techniques (such as huffmann coding, arithmetical coding, dictionary-based methods, predication by partial match and context tree weighting), with CPU instruction specific transformations such as jump instruction modification. [1, 2] Compression of executable code is mainly used for lowering bandwidth usage on transfer and decreasing storage needs in embedded devices.

The compiled code can be split into so called basic blocks, sequences of instructions ending with a single control transfer instruction. The internal program representation used to facilitates our program transformations is the Control Flow Graph. This graph contains basic blocks as nodes, and potential control flow paths as edges.

In my work I show, that reordering instructions using data flow constraints can improve code compression, without changing the original behavior of the code. Using the data usage obtained from the analysis of the disassambled code, a basic, local data flow graph can be produced. A basic block instruction data flow is shown on figure 3. , each instruction is in a separated box, the original position is the code is the first line, the actual instruction is in the second line. The arrows represents data flow within the basic block, indirect dependencies are hidden.

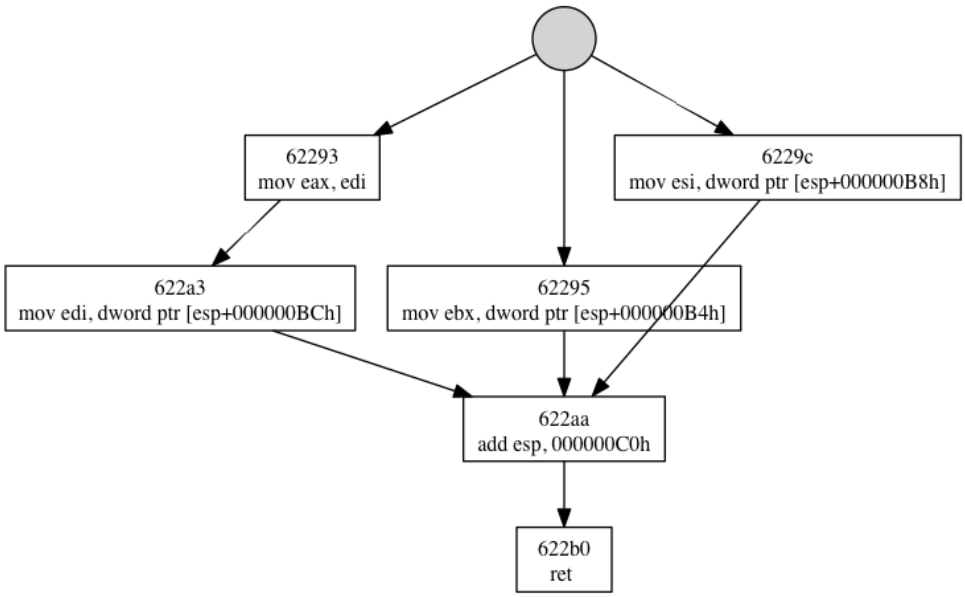


Figure 3: Data dependency graph in a basic block

The order of these instructions can be different from the original, and still the reordered code will be equivalent with the original (in the state machine sense).

In my work I distinguish two kind of data affection (read, write), and use 27 data types include registers: 16 basic x86 registers, 11 eflags, and memory data. The memory data refers to

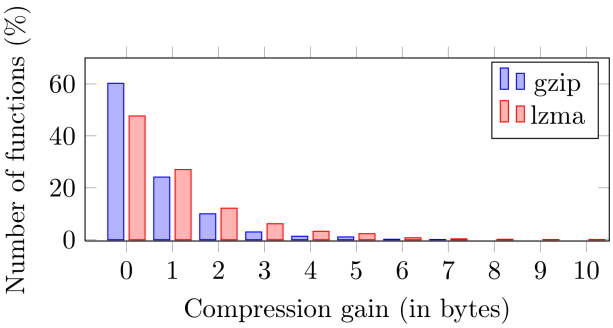


Figure 4: Compressed code size change

any kind of memory read/write including stack instructions. The generated control flow graph has a control flow instruction at the end of each basic block, these instructions are condition and unconditional jump and call instructions are treated as if they write every data type. Among x86 instructions only control flow instructions have relative to current address pointers, that is why reordering instructions can be done by simply changing the instruction's order. The control flow instructions always remain at the end of each basic block.

The complexity of a function is a product of all basic block instruction count within a function. The code generates all possible reordering using depth first search, dependency rules are checked on each search iteration. When all the possible reordering are generated, two different data compression methods are evaluated.

In my work I use the latest stable gzip and lzma software to compress the produced code. Repeating the reordering and compression test on every function of a binary file, a more compressible code can be produced, without modifying the data flow. Using gzip, the compressed code for an avarage function is 0.71% smaller than with the original instruction ordering. Results are even better using lzma, the avarage gain is 1.13%.

A detailed statistics on the compressed code size change can be seen on figure 4. The conclusion is that, more than 40% of functions can have a better compressable reordering than the original one.

References

[1] Árpád Beszédes, Rudolf Ferenc, Tibor Gyimóthy, André Dolenc, and Konsta Karsisto. Survey of code-size reduction methods. *ACM Comput. Surv.*, 35(3):223 - 267, September 2003.

[2] Wenrui Dai, Hongkai Xiong, and Li Song. On non-sequential context modeling with application to executable data compression. *In Data Compression Conference*, 2008. DCC 2008, pages 172 - 181, march 2008.

Non-destructive Examination of Tree Trunks with the Help of National Instruments Devices

István Póser

The parameters of vibration signals spreading in a special type of tree trunk change depending on the damaged or dead portions inside it, because the vibrations will only spread in the higher density regions. So the time of the signal flight, the amplitude and other parameters of the signals will change too depending on the inner structure of the tree trunk. I'm measuring these parameters, primarily the time of flight and the amplitude with piezolelctrical accelerometers and a prototype measurement system based on National Instruments devices. The software controlling the measurement system and preprocessing the data is based on LabVIEW. I'm trying to build up a picture of the inner structure of the tree trunk from multiply generation and measurement cycles on a PC. This first approach is required to determine the parameters that should be measured and algorithms that should be used in the future. The final goal is to build a portable device capable of examining the inner structure of a tree trunk on the field without the need for using a computer.

The purpose of this paper is to show the different type of National Instruments devices, that I used, and the results, that I achieved in my work so far.

Acknowledgements

This work was supported by Gabor Nagy and Laszlo Toth.

Workflow processing using SAP Objects

Attila Selmeci, Tamás Gábor Orosz

The SAP, the developer of the leading Enterprise Resources Planning System in the World, has been providing business solutions in different areas required by the market for more than 40 years. The basic development environment of the SAP applications is smoothly implemented into the system. All business applications are developed in the internal ABAP (Advanced Business Application Programming) language. This language has its origin from the old style COBOL, PL1, but in the last 40 years the company had enough time to implement features from new technologies, programming languages. In the last ten years SAP implemented and uses for its own applications the object oriented programming paradigm, which is an extension of the original ABAP.

The business logic is implemented as reports, so-called transactions and function modules. Reports generate lists, transactions help to store, change business data in the system through many input and output screens defining a single logical unit of work. The function modules are special modularization units executing a determined functionality (like booking). Grouped function modules can share data through the main program of the group, which cannot be touched directly only via the connected function modules. This approach is similar to the static methods and attributes of an OOP class.

The Business Framework Architecture provides a special environment for internal and external usage of the system services by collecting them into Business Components and offering the real business entities as so-called Business Objects. The Business Objects (a determined entity from an entity type) are instances of the Business Object types. This Business Object layer extends the original report, transaction and function module layer with an object-oriented view, which is usable from outside as well. The Business Object Types contain as subcomponents attributes, methods, and events as well. There are special attributes, the key fields, which refer to the real, underlying data model. The other attributes can be simple fields, structures, arrays or even references to other Business Object Types. The events are information to the "world" about status changes of the Business Object. For example an employee can be hired or fired, an invoice can be approved, rejected, booked, parked as well. If a status change would be interesting for the system it can be publicized by an event. The methods execute the status changes of the business objects, like book, approve for an invoice. A method (as an attribute) can be instance specific or static (class level), and it can have parameters as well. There are special methods in the Business Object world, like CreateFromData, GetList, etc. These can get information from the available entities of the entity type or create a new instance, entity from scratch in the database. There is a unique meaning of public methods, because it means released for external availability. Only these methods can be called from outside of the system. These public methods of Business Objects are the BAPIs (Business Application Programming Interfaces).

The main constituent of process control is an object, which can store its current state; the state can be modified from outside, and the status changes can be publicized. The Business Objects are good candidate to be process control components.

In our study we uncover the options to build up process flows or workflows within SAP systems, and check the potential of creating outside driven workflows. We try to stretch the borders of the study to show the possibilities of creating and using system-wide workflows or processes by help of Business Objects.

Acknowledgements

This work was supported by TAMOP-4.2.2/B-10/1-2010-0012 Programme.

AST-based source code analysis and visualization

Melinda Simon

Developing complex software systems is a long-term, difficult process which requires a lot of planning. We have to guarantee the quality of the system, which means in this case that the software works as it is expected and satisfies the needs of the procurers. To lower the costs of testing and avoiding later bugfixes, the system is continously analyzed by automated testing during the development from the very beginning. These tests can be unit tests, which can validate and/or verify the implementation of single unit (e.g. a class or a function) or integration test cases, which investigates the cooperation between different components of the system [2].

Although we have some good tools to use when writing unit tests (JUnit [3] and other JUnit-based tools), the data used for testing have to be chosen manually. To ease our work we can test and visualize the code coverage of our unit test by various open-source frameworks, such as Coverclipse or CodeCover and a lot more, which shows how much of a software's source code has been tested and can point to the weaknesses of testing in our project.

Similarly to unit tests, we would like to see the code coverage when running integration test to highlight the riskious parts of the project. There are several tools to run integration tests and some of these can count information on integration tests' code coverage, e.g. Sonar project [4]. It is not enough to have a good code coverage by integration tests if we have a poor result from unit tests' code coverage. Having code coverage for both unit tests and integration tests, we have to combine these informations and use a new metrics to indicate the quality of the developed code and it's test suites. For this purpose I created a simple Eclipse-plugin, which analyzes the Abstract Syntax Tree (AST) [1], counts the code coverage from calls started from a different component of the system. This plugin can visualize the colored AST and print a report containing various information about the code coverage when a call comes from a different module, e.g. which methods were involved, what parts of these methods were traversed how many times, and so on. Know the weak points of our integration tests we can improve our test cases and grant the quality of the software more effectively.

References

[1] Aho, A. V., Lam, M. S., Sethi, R., Ullman, J. D.: "Compilers: Principles, Techniques, and Tools", 2nd Edition, Addison-Wesley (2006)

[2] Perry, W. E.: "Effective Methods for Software Testing", 2nd Edition, John Wiley and Sons, (2000)

[3] JUnit.org Resources for Test Driven Development, `http://junit.org`

[4] Sonar Project, `http://www.sonarsource.org`

Extracting Geometrical Features of Discrete Images from Their Projections

Tamás Sámuel Tasi, Péter Balázs

The task of *Reconstruction Tomography* is to produce 2-dimensional cross-section images of an otherwise 3-dimensional target object from several projections, which would reveal the inner structure of the object under investigation. Due to the nature of the problem, any prior information about the build-up or conformation of the test object is more than welcome to assist the task, usually leading to a reconstruction of better quality. This especially applies for *Discrete Tomography*, designated by DT in short, where the number of projections to work with is very limited. In return, we assume that the image only holds a few discrete intensity values known beforehand. Nevertheless, in order to obtain appropriate results we have to limit the search space by exploiting the additional information about the object given prior to reconstruction.

Most of the reconstruction algorithms in DT rely heavily on such 'a priori' knowledge. Extraction of these features could be attempted from the projections themselves, e.g. by using the wide range of tools of Artificial Intelligence. We succesfully used learning algorithms, like Feedforward Neural Networks to retrieve geometrical properties of images, such as convexity and connectivity [1], as well as estimating the perimeter of the shape [2] and the number of distinct intensities present in the image [3] from the projection data only. Knowledge gained during this preliminary step could then be used to ease the task of choosing the proper reconstruction method, or setting its parameters appropriately.

The work of the second author was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences and by the OTKA PD100950 project of the Hungarian Scientific Research Fund.

References

[1] Gara, M., Tasi, T. S. and Balázs, P.: Learning Connectedness and Convexity of Binary Images from Their Projections. *Pure Mathematics and Applications*, 20(1–2):27–48, 2009.

[2] Tasi, T. S., Hegedűs, M. and Balázs, P.: Perimeter Estimation of Some Discrete Sets from Horizontal and Vertical Projections. *The Ninth IASTED International Conference on Signal Processing, Pattern Recognition and Applications*, in press, 2012.

[3] Gara, M., Balázs, P. and Tasi, T. S.: Machine Learning as a Preprocessing Phase in Discrete Tomography. *Applications of Digital Geometry and Mathematical Morphology, Lecture Notes in Computer Science*, in press, 2012.

Ideas for Improving Efficiency of Procedural Abstraction

Tamás Szirbucz

Memory consumption in embedded systems is a critical arrea. The smaller code size, the less memory is needed on the device or more features can be implemented in the system. An effective method for reducing code size is the code factoring. An algorithm called procedural abstraction has been previously implemented in GCC [1]. It finds identical copies of code sequences in the program and abstract them out into functions, replacing the original occurences with function calls.

Currently we are investigating how the original algorithm can be made more effective. Below, I present some ideas we are currently focusing on. The original algorithm works on basic blocks. However bigger single-entry/single-exit sections of code can turn out to be clones. Finding such duplicates is computationally more expensive, but can give higher gain as well.

It should also be improved with variable renaming. In many cases two sequences differ in just variable names, that the current algorithm doesn't handle. Otherwise variable renaming increases the overhead of abstractions, and also has higher computing cost.

The current choosing strategy of abstractable sequences is a greedy algorithm. It choose the sequences with higher gain first. I will show, it is less optimal than abstracting the longest sequences first.

At the conference, beside presenting the above ideas, I will discuss implementation problems of these modifications.

References

[1] Gábor Lóki, Ákos Kiss, Judit Jász, and Árpád Beszédes. *Code Factoring in GCC* In Proceedings of the 2004 GCC & GNU Toolchain Developers' Summit (GCC 2004), Ottawa, Ontario, Canada, June 2-4, 2004.

Efficiency and Accuracy Analysis of a Fuzzy Single-Stroke Character Recognizer with Various Rectangle Fuzzy Grids

Alex Tormási and László T. Kóczy

Nowadays the market share of portable computer devices is continuously growing. In some cases the use of physical keyboards in portable devices is not possible at all. This is why designers are in need of new methods for data input, such as software-based virtual keyboards combined with touch interfaces, further voice and handwriting recognition.

Processing written text by computers nevertheless has a long history. In this field there are still many ongoing research and development projects aiming to achieve the more accurate recognition of handwriting. In her study LaLomia [1] determined 97% as the general user acceptance rate for handwriting recognizers.

In this paper the results of a study on the efficiency and accuracy of a fuzzy logic-based [2] single-stroke character recognizer are presented by using various rectangle fuzzy grid cells in the feature extraction phase.

During the design of the recognition algorithm limiting the resources needed for the method was targeted as a primary goal besides the acceptable recognition accuracy. A solution has been worked out to eliminate geometrical transformations from the method so we could reduce the overall computational complexity. In the early stage of the development we decided to handle the stroke segmentation as a separate problem so we could focus on the concept of the recognition engine.

Each stroke could be represented as a continuous function sampled by the digitizer tablet. The system collects all the coordinates in chronological order representing the pen movement. Due to hardware bottlenecks the input device is not capable of collecting all parts of the signal and the distance between the sampled points may differ.

During the pre-processing phase the input stroke is re-sampled. The first and last points of the stroke are stored for reference. After that, the filtering algorithm calculates the distance between the points. If the distance reaches the minimum threshold, the point will be added to the re-sampled stroke if however it is not it will be filtered out.

Input signals are identified by the width/height ratio of the stroke and the average number of stroke-points in the rows and columns of the grid. Our tests pointed out a big disadvantage of crisp grids (grids with sharp borders). If the angular offset of the input stroke and the etalon symbol are different then the distribution of the points in the grid will also differ. This might cause a high reduction in recognition rates. As a solution we designed a grid with blurred boundaries which will be referred to as fuzzy grid (inspired by fuzzy logic).

The symbol set is a modified version of Palm's Graffiti single-stroke alphabet and it contains 26 different symbols. The rule base was determined by a subset of the collected samples. The extracted stroke-features are used as input parameters for the fuzzy rules [3, 4]. The number of rules is equivalent to the number of the symbols in the base set. The consequent part of the rules represents the degree of matching of the parameters of the input stroke and of the parameters of the symbol represented by the given rules. The best fitting rule will be chosen as the output of the inference.

Our first implemented system [5] used a four-by-four fuzzy grid and reached 94.3% accuracy. Our more recent work compared the efficiency of the system with different N-by-N (square) fuzzy grids using trapezoidal membership functions of rows and columns defined to cover each point constituting Ruspini-partitions [6]. As a result 99.4% recognition rate was achieved with the 6-by-6 fuzzy grid. In this work different numbers of rows and columns were analyzed for the N-by-M fuzzy grids in order to find its optimal size. The new system achieved 99.2% recognition rate with a 6-by-4 fuzzy grid which almost reached the best accuracy of the square fuzzy grid, however it resulted in a decreased computational complexity.

To make the system more accurate the algorithm includes a training phase where the specific features of the user's handwriting is learned but this function was disabled during the tests to avoid its effects on the accuracy. In the adaptation phase an evolutionary algorithm [7] was used due to the limited available hardware resources. The results clearly showed that bacterial evolutionary algorithms [8, 9] were suitable for tuning fuzzy sets. During the adaptation the system had to consider the features of the previously stored samples as much as the new input stroke. All parameters of the new symbol had to fit to the tuned fuzzy set of the target symbol as much as possible without decreasing the fitness of the stored samples. At the same time the method had to minimize the overlap of the target and non-target fuzzy rules without modifying the fitness of the samples stored in non-target symbol classes. Classical evolutionary solutions could not be used due to the special constraints in which the different dynamic fitness functions had to be deployed for the different symbols at the same process. Without these constraints the fuzzy sets would overlap in the different rules which would decrease the recognition rate. The bacterial evolutionary algorithm has been extended with "punish" and "reward" option. The modified algorithm solved the overlap-problem and resulted in an increased average recognition rate.

Acknowledgements

This paper was supported by the National Scientific Research Fund Grant OTKA K75711, a Széchenyi István University Main Research Direction Grant and EU grant TÁMOP 421 B.

References

[1] LaLomia, M.J.: User Acceptance of Handwritten Recognition Accuracy. *In: Companion Proc. CHI '94*, p. 107. New York (1994)

[2] Zadeh, L.A.: Fuzzy sets. *In: Inf. Control*, 8:338-353 (1965)

[3] Mamdani, E.H., Assilian, S.: An Experiment in Linguistic Synthesis with a Fuzzy Logic Controller. *In: International Journal of Man-Machine Studies* Vol. 7. pp. 1-13 (1975)

[4] Takagi, T., Sugeno, M.: Fuzzy identification of systems and its applications to modeling and control. *In: IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-15, pp. 116-132 (1985)

[5] Tormási, A., Botzheim, J.: Single-stroke Character Recognition with Fuzzy Method. *In: Balas, V.E. et al. (Eds.): New Concepts and Applications in Soft Computing*, SCI 417, pp. 27-46. Springer-Verlag, Berlin Heidelberg (2012)

[6] Ruspini, E.H.: A new approach to clustering. *In: Information Control*, 15(1):22-32 (1969)

[7] Holland, J.H.: Adaption in Natural and Artificial Systems. The MIT Press, Cambridge, Massachusetts (1992)

[8] Nawa, N.E., Furuhashi, T.: Fuzzy system parameters discovery by bacterial evolutionary algorithm. *In: IEEE Transactions on Fuzzy Systems*, 7(5): 608-616 (1999)

[9] Cabrita, C., Botzheim, J., Ruano, A.E. and Kóczy, T.L.: Genetic programming and bacterial algorithm for neural networks and fuzzy systems design. *In: International Conference on Intelligent Control Systems and Signal Processing, (ICONS IFAC)*. Faro, Portugal, pp. 500-505 (2003)

Geometric Newton-Raphson Methods for Plane Curves

Gábor Valasek, Levente Sallai, András Jámbori, Júlia Horváth

Plane curves play an important role in computer graphics, geometric modeling, GIS systems, and other applications. Queries on these curves, such as findig the closest point of a curve to a given point of the plane, or finding the intersection of two plane curves, often require the solution of a system of non-linear equations or finding the roots of a non-linear function. The Newton-Raphson method is a popular choice for solving these problems iteratively.

Kallay presented a geometric Newton-Raphson strategy in [1]. Each new guess was computed by solving a geometric approximation of the problem at the current guess. The geometric approximation consisted of substituting the curve with its osculating circle, and using the solution of the query on the osculating circle to choose the new guess.

Our paper presents a secant method-like modification of Kallay's algorithm and compare it with the classic and Kallay's geometric Newton-Raphson methods in point-curve distance and curve-curve intersection queries, with emphasis on computational cost and robustness issues. The osculating circle is approximated by fitting a circle to points on the curve, avoiding the evaluation of the curve's derivatives. Another modification of Kallay's method, that uses osculating parabolas and their approximations instead of osculating circles, is also presented and evaluated.

Acknowledgements

This research was supported by the project TÁMOP-4.2.1/B-09/1/KMR-2010-003 of Eötvös Loránd University.

References

[1] Kallay M. A geometric Newton-Raphson strategy. Computer Aided Geometric Design, Volume 18, Number 8, October 2001 , pp. 797-803(7)

A Novel Optimization-Based Reconstruction Algorithm for Multivalued Discrete Tomography

László Varga, Péter Balázs and Antal Nagy

Transmission tomography [1, 2] is the reconstruction of objects from their projections. This is usually done by exposing the object of study to some type of radiation, and measuring the transmitted energy on the other side. The object will absorb some fraction of the radiation and from this, one can derive the summed density of the object along the paths of the beams. By taking such projections from many different directions it is possible to reconstruct the inner structure of the object.

In discrete tomography [3, 4] one also assumes that the object of study consists of only few known materials. With this information it is possible to reduce the number of required projections significantly. This can be useful in practical applications since the projection acquisition can be cost-, or time-consuming, or in some cases the high amount of radiation caused by taking many projections can damage the object [5].

We developed a new reconstruction algorithm that can provide accurate reconstructions of objects in the binary and non-binary case of discrete tomography, by minimizing an energy function with a novel optimization process.

We also tested the algorithm by comparing it to other reconstruction methods, in a series of software test.

Acknowledgements

This research was in part supported by the TÁMOP-4.2.1/B-09/1/KONV-2010-0005 project of the Hungarian National Development Agency co-financed by the European Union and the European Regional Development Fund. The work of the second author was also supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences and the PD100950 grant of the Hungarian Scientific Research Fund (OTKA).

References

[1] A. C. Kak, M. Slaney, Principles of Computerized Tomographic Imaging, *IEEE Press*, New York, 1999.

[2] G.T. Herman, Fundamentals of Computerized Tomography, Image Reconstruction from Projections, 2nd edition, *Springer-Verlag*, London, 2009.

[3] G.T. Herman, A. Kuba (Eds.), Discrete Tomography: Foundations, Algorithms and Applications, *Birkhäuser*, Boston, 1999.

[4] G.T. Herman, A. Kuba (Eds.), Advances in Discrete Tomography and Its Applications, *Birkhäuser*, Boston, 2007.

[5] S. van Aert, K.J. Batenburg, M.D. Rossell, R. Erni, G. Van Tendeloo. Three-dimensional atomic imaging of crystalline nanoparticles, *Nature* 470, pp. 374-377 (2011).

List of Participants

Antyipin, Artyom: Eötvös Loránd University, Faculty of Informatics, H-1117 Budapest, Pázmány Péter sétány 1/C, E-mail: artyom@pnyf.inf.elte.hu

Árgilán, Viktor: University of Szeged, Gyula Juhász Faculty of Education, 6. Boldogasszony Ave, Szeged H-6725, E-mail: gilan@jgypk.u-szeged.hu

Bodnár, Péter: University of Szeged, Institute of Informatics, H-6720 Szeged, Árpád tér 2., E-mail: bodnaar@inf.u-szeged.hu

Borsi, Zsolt: Eötvös Loránd University, Faculty of Informatics, H-1117 Budapest, Pázmány Péter sétány 1/C, E-mail: bzsr@inf.elte.hu

Bóta, András: University of Szeged, Institute of Informatics, H-6720 Szeged, Árpád tér 2., E-mail: bandras@inf.u-szeged.hu

Dávid, Balázs: University of Szeged, Institute of Informatics, H-6720 Szeged, Árpád tér 2., E-mail: davidb@inf.u-szeged.hu

Divéki, Gabriella: Subotica Tech - College of Applied Sciences, Informatics, Marka Oreskovica 16, 24000 Subotica, Serbia, E-mail: diveki.gabriella@gmail.com

Dobján, Tibor: College of Dunaujvaros, Hungarian Acoustic and Industrial Diagnostic Testing Laboratory, H-2400 Dunaujvaros, Tancsics Mihaly street 1/A, E-mail: dobjan.tibor@mail.duf.hu

Fazekas, Ádám: Faculty of Informatics, University of Debrecen, H-4028 Debrecen, Kassai út 26., E-mail: fazekas.adam@inf.unideb.hu

Fodor, Bálint: Budapesti Műszaki és Gazdaságtudományi Egyetem, Automatizálási és Alkalmazott Informatikai Tanszék, 1111 Budapest, Műegyetem rakpart 3-9., E-mail: balint.fodor@aut.bme.hu

Földi, Edit: Faculty of Informatics, University of Debrecen, H-4028 Debrecen, Kassai út 26., E-mail: edit.foldi@gmail.com

Fülöp, András: Faculty of Informatics, University of Debrecen, H-4028 Debrecen, Kassai út 26., E-mail: fulibacsi@gmail.com

Gara, Mihály: University of Szeged, Institute of Informatics, H-6720 Szeged, Árpád tér 2., E-mail: gara@inf.u-szeged.hu

Gonda, László: Faculty of Informatics, University of Debrecen, H-4028 Debrecen, Kassai út 26., E-mail: golaczy@gmail.com

Griechisch, Erika: University of Szeged, Institute of Informatics, H-6720 Szeged, Árpád tér 2., E-mail: grerika@inf.u-szeged.hu

Gulyás, Gergely: University of Szeged, Institute of Informatics, H-6720 Szeged, Árpád tér 2., E-mail: gulyasg@inf.u-szeged.hu

Hantos, Norbert: University of Szeged, Institute of Informatics, H-6720 Szeged, Árpád tér 2., E-mail: nhantos@inf.u-szeged.hu

Hegedüs, László: Faculty of Informatics, University of Debrecen, H-4028 Debrecen, Kassai út 26., E-mail: hegedus.laszlo@inf.unideb.hu

Hegedűs, Péter: University of Szeged, Institute of Informatics, H-6720 Szeged, Árpád tér 2., E-mail: hpeter@inf.u-szeged.hu

Hodován, Renáta: University of Szeged, Institute of Informatics, H-6720 Szeged, Árpád tér 2., E-mail: hodovan@inf.u-szeged.hu

Horváth, Júlia: Eötvös Loránd University, Faculty of Informatics, H-1117 Budapest, Pázmány Péter sétány 1/C, E-mail: horvath_julia@citromail.hu

Jámbori, András: Eötvös Loránd University, Faculty of Informatics, H-1117 Budapest, Pázmány Péter sétány 1/C, E-mail: andras.jambori@gmail.com

Keresztury, Balázs: Corvinus University of Budapest, Department of Computer Science, H-1093 Budapest, Fővám tér 8., E-mail: balazs.keresztury@uni-corvinus.hu

Kovács, Péter: Eötvös Loránd University, Faculty of Informatics, H-1117 Budapest, Pázmány Péter sétány 1/C, E-mail: kovimesterr@gmail.com

Lázár, Edit: Faculty of Informatics, University of Debrecen, H-4028 Debrecen, Kassai út 26., E-mail: ledit4@gmail.com

Lefkovits, Szidónia: Petru Maior University Tirgu Mures, Mathematics-Computer Science, Nicolae Iorga 1, Tirgu Mures, E-mail: szidonia.lefkovits@science.upm.ro

Lévai, Balázs László: University of Szeged, Institute of Informatics, H-6720 Szeged, Árpád tér 2., E-mail: Levai.Balazs.Laszlo@stud.u-szeged.hu

Lócsi, Levente: Eötvös Loránd University, Faculty of Informatics, H-1117 Budapest, Pázmány Péter sétány 1/C, E-mail: locsi@inf.elte.hu

Mosavi, Amir: Faculty of Informatics, University of Debrecen, H-4028 Debrecen, Kassai út 26., E-mail: amir.allen.hu@gmail.com

Móra, György: University of Szeged, Institute of Informatics, H-6720 Szeged, Árpád tér 2., E-mail: gymora@inf.u-szeged.hu

Nagy, István: University of Szeged, Institute of Informatics, H-6720 Szeged, Árpád tér 2., E-mail: nistvan@inf.u-szeged.hu

Németh, József: University of Szeged, Institute of Informatics, H-6720 Szeged, Árpád tér 2., E-mail: nemjozs@inf.u-szeged.hu

Orbán, György: Eötvös Loránd University, Faculty of Informatics, H-1117 Budapest, Pázmány Péter sétány 1/C, E-mail: o.gyorgy@gmail.com

Orosz, István: Óbuda University, AREK, H-8000 Székesfehérvár, Kassai utca 30, E-mail: orosz.istvan@arek.uni-obuda.hu

Ozsvár, Zoltán: University of Szeged, Institute of Informatics, H-6720 Szeged, Árpád tér 2., E-mail: ozsvar.zoltan@gmail.com

Paróczi, Zsombor: BME, TMIT, 1117, Budapest, Magyar tudósok körútja 2., I.L.106a, E-mail: paroczi@tmit.bme.hu

Pándi, Krisztián: Budapest University of Technology and Economics Budapest, Hungary, Magyar tudósok krt. 2., Budapest 1117, E-mail: pandi.krisztian@aut.bme.hu

Póser, István: Faculty of Informatics, University of Debrecen, H-4028 Debrecen, Kassai út 26., E-mail: poser.istvan@gmail.com

Selmeci, Attila: Óbudai Egyetem, H-8000 Székesfehérvár, Budai út. 45., E-mail: selmeci.attila@arek.uni-obuda.hu

Simon, Melinda: Eötvös Loránd University, Faculty of Informatics, H-1117 Budapest, Pázmány Péter sétány 1/C, E-mail: melinda@elte.hu

Szirbucz, Tamás: University of Szeged, Institute of Informatics, H-6720 Szeged, Árpád tér 2., E-mail: szirbucz@inf.u-szeged.hu

Szügyi, Zalán: Eötvös Loránd University, Faculty of Informatics, H-1117 Budapest, Pázmány Péter sétány 1/C, E-mail: lupin@ludens.elte.hu

Tasi, Tamás Sámuel: University of Szeged, Institute of Informatics, H-6720 Szeged, Árpád tér 2., E-mail: ttasi@inf.u-szeged.hu

Tormási, Alex: Széchenyi István University, Department of Automation, Egyetem tér 1, Győr, H-9026, E-mail: tormasi@sze.hu

Valasek, Gábor: Eötvös Loránd University, Faculty of Informatics, H-1117 Budapest, Pázmány Péter sétány 1/C, E-mail: valasek@inf.elte.hu

Varga, László: University of Szeged, Institute of Informatics, H-6720 Szeged, Árpád tér 2., E-mail: vargalg@inf.u-szeged.hu

Notes

Supported by TÁMOP-4.2.2/B-10/1-2010-0012 Programme for Broadening the knowledge base and supporting the long term professional sustainability of the Research University Centre of Excellence at the University of Szeged by ensuring the rising generation of excellent scientists.